

AN INTEGRATED NETWORK OF ESTROGEN RECEPTORS ALPHA AND BETA,  
AND COREGULATORS, FOR DECIPHERING ESTROGEN SIGNALING  
IN BREAST CANCER CELLS

BY

TZE HOWE CHARN

DISSERTATION

Submitted in partial fulfillment of the requirements  
for the degree of Doctor of Philosophy in Bioengineering  
in the Graduate College of the  
University of Illinois at Urbana-Champaign, 2010

Urbana, Illinois

Doctoral Committee:

Professor John A. Katzenellenbogen, Chair  
Professor Benita S. Katzenellenbogen  
Assistant Professor Sheng Zhong  
Assistant Professor Yingxiao Wang

## ABSTRACT

The nuclear hormone receptors, ER $\alpha$  and ER $\beta$ , are known to regulate the transcriptional response programs of their target cells, including breast cancer cells. However, their comparative abilities to localize at chromatin binding sites across the genome, the recruitment of major coregulators such as SRC3 and RIP140 by the ERs, and the association of ERs with other transcription factors and chromatin remodeling factors is incompletely understood. Therefore, in this report, we have used both chromatin immunoprecipitation (ChIP) on microarray (ChIP-chip) and ChIP sequencing (ChIP-seq) approaches in breast cancer cells containing three different complements of ERs (ER $\alpha$  alone, ER $\beta$  alone, or ER $\alpha$  + ER $\beta$ ) treated with estradiol to define the cartography of chromatin binding sites for ER $\alpha$ , ER $\beta$ , and the coregulators SRC3 and RIP140. We found that ER $\alpha$  and ER $\beta$  bind to a similar, large number of sites in breast cancer cells containing only one ER subtype, but the two ERs appear to restrict each others chromatin binding and occupy fewer sites in cells containing both ER $\alpha$  and ER $\beta$ . We also observed that there are differences in terms of enriched motifs in ER $\alpha$  and ER $\beta$  binding sites, with ER $\alpha$  binding sites enriched in GATA and FOXA1 motifs, but ER $\beta$  sites being preferentially enriched in E2F motifs. In addition, in cells containing both ER $\alpha$  and ER $\beta$ , ER $\alpha$  appears to displace ER $\beta$  so that ER $\beta$  binds to sites substantially less enriched in estrogen response element (ERE) sequence motifs.

Gene chip microarray transcriptional profiling and gene ontology analysis delineated a core set of genes that correlate with ER $\alpha$  proliferative and ER $\beta$  anti-proliferative effects in breast cancer cells. ER $\beta$  activation by estradiol was associated with the inhibition of genes associated with cell proliferation and the up-regulation of pro-apoptotic genes and genes responding to DNA damage, whereas ER $\alpha$  activation was associated with the downregulation of pro-apoptotic

genes and genes repressing transcription. Analysis of chromatin binding of SRC3 and RIP140 by ChIP-seq revealed that these coregulators are recruited preferentially to ER binding sites of estrogen-induced genes, whereas they are seldom recruited to ER binding sites of hormone-repressed genes, indicating that the SRC3-RIP140 complex is likely to be playing a central role in the induction of ER targeted genes. Our findings suggest an integrated model in which the actions of cofactors such as FOXA1, GATA3, and E2F enforce the selectivity and range of ER $\alpha$  and ER $\beta$  binding and gene regulatory actions, with the coregulators SRC3 and RIP140 preferentially supporting the stimulatory actions of both receptors on gene expression.

## **ACKNOWLEDGMENTS**

I would like to express my sincerest gratitude to my advisers, Drs. John and Benita Katzenellenbogen and Dr. Edison Tak-Bun Liu for all their support and guidance over the years; without them, an electrical engineer like me will never be able to juggle life science research. I would also like to thank my parents, Cheng Chong Charn and Yoke Meng Ho for their constant encouragement and emotional support. I am fortunate to have been able to work in two wonderful groups - the Katzenellenbogen and Liu Groups where I've made many friends. I would especially like to thank Dr. Daniel Barnett, Dr. Edmund Chang, Dr. Joshua Stender, Sung-Hee Park, and Vega Vinsensius B. for all their wonderful help and lively discussions over the years.

My deepest thanks to my committee advisers and the faculty members who have given me their time and helpful advice over the years, Drs. Sheng Zhong, Peter Yingxiao Wang and Michael Insana. Finally thank you to the staff of the Department of Bioengineering at the University of Illinois at Urbana-Champaign and the Genome Institute of Singapore community for all their support and friendly advice.

## TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION .....	1
1.1 Estrogen Receptors (ERs) .....	1
1.2 Molecular Mechanisms of Estrogen Action .....	2
1.3 Roles of ER $\alpha$ and ER $\beta$ in Breast Cancer .....	2
1.4 Expression Profiling of ER Action .....	3
1.5 Genome-Wide Analysis of ER $\alpha$ and ER $\beta$ Binding Sites .....	5
1.6 Recruitment of Coregulators to ER Binding Sites .....	6
1.7 Aims of the Thesis Research .....	7
1.8 References .....	9
CHAPTER 2: EXPERIMENTAL METHODS AND VALIDATION .....	12
2.1 Introduction .....	12
2.2 ER Custom-Designed ChIP-Chip Tiling Array .....	14
2.3 ER ChIP-Chip Binding Sites Processing Steps .....	15
2.4 ChIP-Sequencing of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 Binding Sites .....	17
2.5 Peak Calling for ChIP-Sequencing Tags .....	19
2.6 Transcriptional Profiling of ER $\alpha$ and ER $\beta$ Actions .....	20
2.7 Tables and Figures .....	22
2.8 References .....	31
CHAPTER 3: LIGAND REGULATION OF CHROMATIN BINDING SITES OF ESTROGEN RECEPTORS $\alpha$ AND $\beta$ : SUBTYPE COMPETITION, MUTUAL RESTRICTION, AND LIGAND SELECTIVITY .....	32
3.1 Abstract .....	32
3.2 Introduction .....	33
3.3 Materials and Methods .....	35
3.4 Results .....	36
3.5 Discussion .....	47
3.6 Table and Figures .....	51
3.7 References .....	60
CHAPTER 4: A NETWORK OF ESTROGEN RECEPTORS $\alpha$ AND $\beta$ , AND COREGULATORS SRC3 AND RIP140, IN BREAST CANCER TRANSCRIPTIONAL REGULATION .....	62
4.1 Abstract .....	62
4.2 Introduction .....	63
4.3 Materials and Methods .....	65
4.4 Results .....	67
4.5 Discussion .....	80
4.6 Tables and Figures .....	84
4.7 References .....	114
CHAPTER 5: CONCLUSION .....	118

## **CHAPTER 1**

### **INTRODUCTION**

#### **1.1 ESTROGEN RECEPTORS (ERs)**

Estrogens play key roles in many aspects of reproductive physiology, development, and metabolism, and they are also involved in several disease states, including breast and endometrial cancers [1, 2]. The effects of estrogens are mediated through two estrogen receptors, estrogen receptor alpha (ER $\alpha$ ) and beta (ER $\beta$ ), that function as ligand-modulated transcription factors, up- and down-regulating gene expression in a target tissue-selective manner [3, 4]. The presence of ER $\alpha$  in breast cancer cells and in various tissues is associated with enhanced proliferation in response to estrogens, whereas several studies have implicated ER $\beta$  as exerting antiproliferative effects [5-9].

ER $\alpha$  and ER $\beta$  are members of the nuclear receptor superfamily, representing a class of signal-activated DNA-binding transcription factors that respond to small molecule ligands, such as estrogens. Both ER $\alpha$  and ER $\beta$  each contain a ligand-binding domain (LBD), a DNA-binding domain (DBD), and two activation domains (AF-1 and AF-2). ER $\alpha$  and ER $\beta$  are highly homologous in their DNA-binding domains (97% identity), but they are quite different in their ligand-binding domains (56% identity) and transcriptional activation function-1 (18% identity) domains. The differences in their ligand-binding domains allow the two ER subtypes to bind certain ligands with high selectivity, for one or the other ER subtype [10-13]. Thus, in cells containing both ERs, it is possible to activate either receptor selectively by using subtype-selective ligands. Also, due to the differences in the AF-1 domains between the two estrogen receptors, ER $\alpha$  and ER $\beta$  can exert receptor-specific function by interacting with and recruiting different proteins.

## 1.2 MOLECULAR MECHANISMS OF ESTROGEN ACTION

Estrogen acts through its receptors, ER $\alpha$  and ER $\beta$ , to control patterns of gene expression in target tissues. It is now well documented that ER $\alpha$  and ER $\beta$  can exist in cells as homodimers when present alone, and additionally as heterodimers when present together. The ERs dimers are responsible for the regulation of transcriptional activation.

There are several distinct pathways by which estrogen can mediate the biological processes, namely, classic ER action or non-genomic mechanisms. In the classic estrogen actions, ERs function as dimers and bind to regulatory sites on the chromatin where they recruit a variety of coregulators, histone-modifying enzymes, and other factors to up or downregulate the transcription of hundreds of genes that markedly influence cell phenotype. Some of the chromatin binding sites to which the ERs bind are known as estrogen-response elements (EREs). The consensus ERE consists of a 13bp sequence, GGTCAnnnTGACC. However, recent genome-wide studies on ERs chromatin sites [14-17] showed that many of ER binding sites do not contain the consensus ERE motif. In addition to direct DNA binding, ERs can also mediate estrogen-signaling by “tethering” onto other transcription factor complexes such as stimulating protein 1 (SP1) [18], and activating protein 1 (AP1) [19]. Estrogen can also exert its effects through non-genomic mechanisms where it binds to ERs localized on the plasma membrane and initiates extranuclear signaling through activation of kinase cascades [20, 21].

## 1.3 ROLES OF ER $\alpha$ AND ER $\beta$ IN BREAST CANCER

Estrogen has been implicated in the promotion and growth of breast cancer. Although ER $\alpha$  is usually the predominant ER in breast tumors and is currently used as a indicator for the responsiveness to endocrine therapies, most human breast tumors also coexpress both ERs, with

the level of ER $\beta$  covering a broad range [22-24]. It is well documented that ER $\alpha$  in breast cancer cells is associated with enhanced proliferation in response to estrogens; however, the effects of ER $\beta$  in breast cancer are less clear, although *in vitro* studies have indicated that ER $\beta$  can modulate the proliferation and invasion of breast cancer cells [3, 25]. The modulating ability of ER $\beta$  on ER $\alpha$  and its subsequent loss in a majority of breast tumors [26, 27] suggest that ER $\beta$  can be a tumor suppressor gene.

#### 1.4 EXPRESSION PROFILING OF ER ACTION

With the introduction of the DNA microarray platform, it has become feasible to simultaneously query the expression of thousands of genes with great precision. Our lab and other have taken advantage of this technology to comprehensively study the estrogen-regulated gene expression profiles in breast cancer and osteosarcoma cell lines expressing either ER $\alpha$  or ER $\beta$  [5, 9, 28, 29]. These studies have provided us with a system-wide view of ERs actions on its target genes and have also given us an unprecedented view of the transcriptome dynamics in response to E2 and selective estrogen receptor modulators (SERMs). Although both ER $\alpha$  and ER $\beta$  have been shown to be able to heterodimerize when present in the same cell, the impact of heterodimerization on gene regulation is still unclear. Hence, our lab had preformed several gene expression studies aimed at studying the interplay between ER $\alpha$  and ER $\beta$ , and characterizing the role of ER $\beta$  in influencing the transcriptional activity of ER $\alpha$ . To access the modulatory effects of ER $\beta$ , we used adenovirus-mediated gene delivery method to express ER $\beta$  in MCF-7 (ER $\alpha$  only cells) breast cancer cells to create cells that can express both ER $\alpha$  and ER $\beta$  [3]. This allowed us to examine ER $\alpha$ -mediated transcription (in ER $\alpha$  only cells) and the modulating effects of ER $\beta$  on ER $\alpha$ -mediated transcription (in ER $\alpha$  and ER $\beta$  expressing cells). Our studies indicate



that the co-expression of ER $\beta$  with ER $\alpha$  significantly impacted, both in an enhancing and a suppressing fashion, the E2-induced transcriptional response by ER $\alpha$ . Of the genes modulated by ER $\beta$ , the greatest numbers were associated with transcription factors and signal transduction pathways [3]. We observed from our expression data that ER $\beta$  regulated multiple components in the TGF $\beta$ , SDF1, and semaphorin pathways, which may contribute, at least in part, to the anti-proliferative effects observed when ER $\beta$  is present in the cells. In addition, our data identified a subclass of E2-regulated genes that responds to E2 stimulation only in the presence of ER $\beta$ . This raises the possibility that the regulatory regions of these genes contain sequences which bias for the recruitment of ER $\beta$  but not ER $\alpha$ . A second possibility might be due to the different cofactors recruited by either ER $\alpha$  or ER $\beta$  complexes which enable the ER complexes to access different chromatin regions. Our findings from the above study exemplify the complex relationship between ER $\alpha$  and ER $\beta$ . However, the study was not able to give a complete picture of transcription regulation by the ERs as there was no comparison of ER $\beta$ -mediated transcription regulation in ER $\beta$ -only expressing cells and the expression profiling was also done on gene expression arrays (Affymetrix U133A), which contains only half of the current known genes (~36 thousands genes are known currently).

Therefore, in order to completely understand the interplay between ER $\alpha$  and ER $\beta$  in breast cancer cells, new generation of gene expression chips from Affymetrix or Illumina should be used to study ER $\alpha$ - and ER $\beta$ -mediated transcription regulation on cells expressing ER $\alpha$ -alone, ER $\beta$ -alone, and ER $\alpha$  and ER $\beta$ .

## 1.5 GENOME-WIDE ANALYSIS OF ER $\alpha$ AND ER $\beta$ BINDING SITES

At the chromatin level, ER $\alpha$  and ER $\beta$  interact either via direct binding to regulatory regions encoded with EREs or via indirect binding by tethering to other transcription factors such as AP1, SP1, and NF- $\kappa$ B. Owing to technological limitations, the mechanisms of ERs chromatin binding have been studied only on a small number of endogenous target promoters [30, 31]. However, a number of recent studies have examined the binding of ER $\alpha$  and ER $\beta$  in a less biased way, using the powerful combination of chromatin immunoprecipitation (ChIP) coupled with either DNA microarrays (ChIP-chip) or DNA sequencing (ChIP-seq) [14-17, 32, 33]. These genome-wide ER chromatin binding studies have provided several novel insights into ER chromatin binding events. First, the chromatin sites bound by ER $\alpha$  are located at great distances (often >100 kilobase pairs) from the promoters of the nearest genes. These unexpected locations of ER binding sites relative to gene transcription start sites (TSS) pose a new challenge to studying ER mechanism, as we can no longer assume that ERs occupying binding sites located in the promoter region will be dictating the transcription of the particular gene. Second, long-range activation by liganded ER $\alpha$  bound to distal enhancers was validated by Chromatin Conformation Capture (3C) approaches [34]. The 3C study indicated that ER $\alpha$  can regulate genes at great distances through a looping mechanism. These studies have provided a genome-wide view of the diversity and distribution of ERs binding that was lacking in previous studies. Unfortunately, most of these studies examined only ER $\alpha$  binding, with the exception of three studies done by us and others [17, 32, 33], that examined ER $\beta$  binding sites by ChIP-chip. However, the examination of ER $\beta$  binding sites by the three studies was not as thorough as the studies done on ER $\alpha$ , as either only selected sites in the genome were used (custom-designed arrays [32], only selected chromosomes were studied [17], or only one cell-type was used [17, 33] (cells

expressing ER $\alpha$  and ER $\beta$ , but no examination of ER $\beta$  binding sites in cells expressing ER $\beta$  alone). Clearly, additional studies are needed to examine the characteristics of ER $\beta$  binding sites throughout the whole genome and to investigate ER $\alpha$  and ER $\beta$  collaboration and/or competition for ER binding sites.

Deciphering the transcription interplay between ER $\alpha$  and ER $\beta$  at chromatin binding sites is critical to the elucidation of the mechanism of estrogen-related gene expression in breast cancer. Interpolating high-throughput gene expression profiling studies with the growing body of genome-wide ERs chromatin binding datasets will enable us to further understand the diverse actions of estrogen in breast cancer growth and offer potential new opportunities for therapeutic intervention.

## 1.6 RECRUITMENT OF COREGULATORS TO ER BINDING SITES

ER-mediated transcription regulation is a complex process involving several steps. Upon activation by its ligands, ER $\alpha$  and ER $\beta$  will undergo a conformation change which enhances ER dimerization and subsequent binding to the regulatory regions of its target genes. Once bound on chromatin, the activated ERs recruit the coregulators to form a multiprotein complex that activates the general transcriptional machinery and increases or reduces the expression of target genes [35, 36]. The coregulators recruited by ERs can be broadly divided into coactivators and corepressors. Coactivators are recruited by ERs to enhance transcriptional activation of target genes, whereas corepressors are recruited by ERs to silence gene activity. Recent evidence indicates that the coregulators assist ERs in regulating transcription by (i) modifying histones and (ii) mediating the interactions with the basal transcription apparatus [37, 38]. The modified chromatin allows other DNA-bound transcription factors to be subsequently recruited.

Consequently, a large array of transcription factors and enzymatic activities converge at the ER binding sites in a temporal manner and set the level of gene activation or repression by ERs. Coactivators such as Swi/Snf, CBP/p300, p160/SRC and corepressors such as N-CoR and SMRT have been shown to be relevant for ERs activity [35]. Crystallographic structures of unliganded and agonist-bound ligand binding domains (AF-2 region) of several nuclear receptors (including ER) revealed that the AF-2 region undergoes a ligand-dependent conformation change which permits the formation of a surface that facilitates coregulator interaction with the nuclear receptors [39].

Given the intimate functional relationships between ER and its coregulators in transcriptional regulation, we can conceivably elucidate ER-mediated transcription regulation by examining ER chromatin binding coupled with the recruitment of coregulators. This will enable us to get a clearer picture of the cartographies of ER chromatin sites. We would be able to separate ER binding sites into sites that can impact transcriptional regulatory control and redundant ER binding sites in the genome. Knowing ER sites that are exerting transcriptional regulatory control will allow us to identify ER target genes.

## 1.7 AIMS OF THE THESIS RESEARCH

Intense efforts have been made in the past decades toward the study of ERs in breast cancer since estrogens are implicated in the development of breast cancer. As transcription factors, the biology of ER $\alpha$  and ER $\beta$  in breast cancer can be studied by characterizing their transcriptional regulatory impact. This can be done by investigating (i) the chromatin localization of ER $\alpha$  and ER $\beta$ , (ii) the co-recruitment of coregulators to ER binding sites, and (iii) the interplay between ER $\alpha$  and ER $\beta$  at the gene expression level. While tangible progress has been made regarding ER

functions, they only permit tentative extrapolation when ERs functions are considered in the genome-wide context. This is because the biochemical and molecular assays routinely used study only one gene at a time. However, in recent years, with the completion of the sequence of the human genome and the advent of high-throughput technologies, the field is aptly poised to explore ERs functions in breast cancer in a genome-wide, unbiased manner. The main objective of this thesis is to comprehensively analyze the genomic targets of ERs actions. We will employ genome-wide approaches such as DNA microarrays for global gene expression characterization, and ChIP-chip and ChIP-seq assays for protein-chromatin interaction.

In Chapter 3, we explore the genome-wide chromatin binding sites of ER $\alpha$  and ER $\beta$  by ChIP-chip. Our ER binding site studies showed that there is substantial overlap in ER $\alpha$  and ER $\beta$  chromatin binding sites when they are present alone in cells, but that many fewer sites are shared when both ERs are present together. Although each ER subtype restricts the binding site occupancy of the other subtype, overall, ER $\alpha$  appears to dominate the binding of ER $\beta$ . We showed that the binding site regions of both ER $\alpha$  and ER $\beta$  are enriched in estrogen response element (ERE) sequence motifs, but when both ERs are present together in the same cell, ER $\beta$  appears to be displaced by ER $\alpha$  so that it binds to sites substantially less enriched in EREs. Chapter 3 highlights the interplay between ER $\alpha$  and ER $\beta$  in selecting and competing for chromatin binding sites in breast cancer cells.

In Chapter 4, we aim to understand ER $\alpha$ - and ER $\beta$ -mediated transcriptional programs in breast cancer by (i) mapping the localization of ER $\alpha$  and ER $\beta$  and the recruitment of coregulators by ERs to the chromatin and (ii) the subsequent gene regulation by ERs and the coregulators. To this end, we employed ChIP-seq technique to map genomic landscape of both ER $\alpha$  and ER $\beta$ , and the coregulators SRC3 and RIP140 in breast cancer cells. Next, gene expression microarray

analyses were carried out to investigate the gene regulatory effects of ER $\alpha$  and ER $\beta$  in breast cancer cells. Finally, by correlating the global cartographies of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 with the results of the gene expression microarray analyses, we present evidence of ERs direct target genes that may contribute to the ER $\alpha$  proliferative and the ER $\beta$  anti-proliferative nature of these receptors in breast cancer cells.

Both Chapter 3 and Chapter 4 highlight the power of genome-wide approaches in examining ERs functions. It allows for the generation of massive, unbiased data that can provide valuable resources for follow-up studies to further our understanding of ERs actions in breast cancer.

## 1.8 REFERENCES

1. Deroo, B.J. and K.S. Korach, *Estrogen receptors and human disease*. J Clin Invest, 2006. **116**(3): p. 561-70.
2. Nilsson, S., et al., *Mechanisms of estrogen action*. Physiol Rev, 2001. **81**(4): p. 1535-65.
3. Chang, E.C., et al., *Impact of estrogen receptor beta on gene networks regulated by estrogen receptor alpha in breast cancer cells*. Endocrinology, 2006. **147**(10): p. 4831-42.
4. Chang, E.C., et al., *Estrogen Receptors alpha and beta as determinants of gene expression: influence of ligand, dose, and chromatin binding*. Mol Endocrinol, 2008. **22**(5): p. 1032-43.
5. Frasor, J., et al., *Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype*. Endocrinology, 2003. **144**(10): p. 4562-74.
6. Paruthiyil, S., et al., *Estrogen receptor beta inhibits human breast cancer cell proliferation and tumor formation by causing a G2 cell cycle arrest*. Cancer Res, 2004. **64**(1): p. 423-8.
7. Strom, A., et al., *Estrogen receptor beta inhibits 17beta-estradiol-stimulated proliferation of the breast cancer cell line T47D*. Proc Natl Acad Sci U S A, 2004. **101**(6): p. 1566-71.
8. Williams, C., et al., *A genome-wide study of the repressive effects of estrogen receptor beta on estrogen receptor alpha signaling in breast cancer cells*. Oncogene, 2008. **27**(7): p. 1019-32.
9. Frasor, J., et al., *Gene expression preferentially regulated by tamoxifen in breast cancer cells and correlations with clinical outcome*. Cancer Res, 2006. **66**(14): p. 7334-40.
10. Kraichely, D.M., et al., *Conformational changes and coactivator recruitment by novel ligands for estrogen receptor-alpha and estrogen receptor-beta: correlations with*

- biological character and distinct differences among SRC coactivator family members. Endocrinology, 2000. 141(10): p. 3534-45.*
11. Malamas, M.S., et al., *Design and synthesis of aryl diphenolic azoles as potent and selective estrogen receptor-beta ligands. J Med Chem, 2004. 47(21): p. 5021-40.*
  12. Meyers, M.J., et al., *Estrogen receptor subtype-selective ligands: asymmetric synthesis and biological evaluation of cis- and trans-5,11-dialkyl- 5,6,11, 12-tetrahydrochrysenes. J Med Chem, 1999. 42(13): p. 2456-68.*
  13. Stauffer, S.R., et al., *Pyrazole ligands: structure-affinity/activity relationships and estrogen receptor-alpha-selective agonists. J Med Chem, 2000. 43(26): p. 4934-47.*
  14. Carroll, J.S., et al., *Genome-wide analysis of estrogen receptor binding sites. Nat Genet, 2006. 38(11): p. 1289-97.*
  15. Kininis, M., et al., *Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters. Mol Cell Biol, 2007. 27(14): p. 5090-104.*
  16. Lin, C.Y., et al., *Whole-genome cartography of estrogen receptor alpha binding sites. PLoS Genet, 2007. 3(6): p. e87.*
  17. Liu, Y., et al., *The genome landscape of ERalpha- and ERbeta-binding DNA regions. Proc Natl Acad Sci U S A, 2008. 105(7): p. 2604-9.*
  18. Saville, B., et al., *Ligand-, cell-, and estrogen receptor subtype (alpha/beta)-dependent activation at GC-rich (Sp1) promoter elements. J Biol Chem, 2000. 275(8): p. 5379-87.*
  19. Umayahara, Y., et al., *Estrogen regulation of the insulin-like growth factor I gene transcription involves an AP-1 enhancer. J Biol Chem, 1994. 269(23): p. 16433-42.*
  20. Madak-Erdogan, Z., et al., *Nuclear and extranuclear pathway inputs in the regulation of global gene expression by estrogen receptors. Mol Endocrinol, 2008. 22(9): p. 2116-27.*
  21. Razandi, M., et al., *Plasma membrane estrogen receptors exist and functions as dimers. Mol Endocrinol, 2004. 18(12): p. 2854-65.*
  22. Paech, K., et al., *Differential ligand activation of estrogen receptors ERalpha and ERbeta at AP1 sites. Science, 1997. 277(5331): p. 1508-10.*
  23. Katzenellenbogen, B.S. and J. Frasor, *Therapeutic targeting in the estrogen receptor hormonal pathway. Semin Oncol, 2004. 31(1 Suppl 3): p. 28-38.*
  24. Kurebayashi, J., et al., *Expression levels of estrogen receptor-alpha, estrogen receptor-beta, coactivators, and corepressors in breast cancer. Clin Cancer Res, 2000. 6(2): p. 512-8.*
  25. Lazennec, G., et al., *ER beta inhibits proliferation and invasion of breast cancer cells. Endocrinology, 2001. 142(9): p. 4120-30.*
  26. Bardin, A., et al., *Loss of ERbeta expression as a common step in estrogen-dependent tumor progression. Endocr Relat Cancer, 2004. 11(3): p. 537-51.*
  27. Skliris, G.P., et al., *Reduced expression of oestrogen receptor beta in invasive breast cancer and its re-expression using DNA methyl transferase inhibitors in a cell line model. J Pathol, 2003. 201(2): p. 213-20.*
  28. Hall, J.M. and K.S. Korach, *Stromal cell-derived factor 1, a novel target of estrogen receptor action, mediates the mitogenic effects of estradiol in ovarian and breast cancer cells. Mol Endocrinol, 2003. 17(5): p. 792-803.*
  29. Hodges, L.C., et al., *Tamoxifen functions as a molecular agonist inducing cell cycle-associated genes in breast cancer cells. Mol Cancer Res, 2003. 1(4): p. 300-11.*

30. Shang, Y., et al., *Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription*. Cell, 2000. **103**(6): p. 843-52.
31. Metivier, R., et al., *Estrogen receptor-alpha directs ordered, cyclical, and combinatorial recruitment of cofactors on a natural target promoter*. Cell, 2003. **115**(6): p. 751-63.
32. Charn, T.H., et al., *Genome-wide dynamics of chromatin binding of estrogen receptors alpha and beta: mutual restriction and competitive site selection*. Mol Endocrinol. **24**(1): p. 47-59.
33. Zhao, C., et al., *Genome-wide mapping of estrogen receptor-beta-binding regions reveals extensive cross-talk with transcription factor activator protein-1*. Cancer Res. **70**(12): p. 5174-83.
34. Carroll, J.S., et al., *Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1*. Cell, 2005. **122**(1): p. 33-43.
35. Glass, C.K. and M.G. Rosenfeld, *The coregulator exchange in transcriptional functions of nuclear receptors*. Genes Dev, 2000. **14**(2): p. 121-41.
36. McKenna, N.J. and B.W. O'Malley, *Combinatorial control of gene expression by nuclear receptors and coregulators*. Cell, 2002. **108**(4): p. 465-74.
37. McKenna, N.J., R.B. Lanz, and B.W. O'Malley, *Nuclear receptor coregulators: cellular and molecular biology*. Endocr Rev, 1999. **20**(3): p. 321-44.
38. Rosenfeld, M.G., V.V. Lunyak, and C.K. Glass, *Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response*. Genes Dev, 2006. **20**(11): p. 1405-28.
39. Moras, D. and H. Gronemeyer, *The nuclear receptor ligand-binding domain: structure and function*. Curr Opin Cell Biol, 1998. **10**(3): p. 384-91.



## **CHAPTER 2**

### **EXPERIMENTAL METHODS AND VALIDATION**

#### **2.1 INTRODUCTION**

Advances in the DNA microarray technology have enable scientists to simultaneously query the expression of thousands of genes. Researchers can now routinely monitor the genome-wide effects of certain treatments, diseases, and developmental stages at the gene expression level. However, the regulation of genes is ultimately determined by the binding of signal-regulated transcription factors to regulatory sites across the genome and the subsequent recruitment of coregulators by the transcription factors to increase or reduce the expression of the target genes. Identifying these DNA regulatory sites for the various transcription factors is a fundamental question in biology as it can help us understand the specific mechanisms connecting transcription factors chromatin binding, followed by coregulators recruitment and subsequent activation or repression of the genes. As ERs are considered master regulators in breast cancer cells and with ERs belonging to the class of signal-activated DNA-binding transcription factors, many studies were conducted to try to map ER binding sites in the genome using the ERE sequence motif [1-3]. Predicting ERs binding sites by these computational studies have limited resolution as the regions queried for ER binding sites are usually near the promoters of genes and therefore are not done on a truly genome-wide manner.

Recently, identification of ERs binding sites directly, in a genome-wide manner, was achieved by combining chromatin immunoprecipitation (ChIP) coupled with either DNA microarrays (ChIP-chip) or DNA sequencing (ChIP-seq) [4-9]. However, most of the studies only examined ER $\alpha$  chromatin binding. Hence, in order to examine the interplay between ER $\alpha$  and ER $\beta$ , we used adenoviral gene delivery to transduce ER $\beta$  expression in ER $\alpha$ -positive MCF-7 cells, both without and with siRNA directed against ER $\alpha$ . These manipulations generated cells

with three complements of ERs, namely cells containing endogenous ER $\alpha$  only, or ER $\alpha$ +ER $\beta$ , or ER $\beta$  only, as previously described [10] (see Figure 2.1). MCF-7 was chosen for our study as the cell-line is extensively studied and all genomic mapping of ER binding sites published to date have used MCF-7 cells. Although the generation of sub-lines from wild-type MCF-7 may alter cellular composition to some degree, they are nevertheless helpful in dissecting ER $\alpha$  and ER $\beta$  cross-modulation activities.

We went on to examine the localization of ER $\alpha$  and ER $\beta$ , when present together or separately, at ER-binding sites in response to different ligand treatments, using ChIP-chip with a custom designed tiling microarray. The design of our ChIP-chip array is detailed in Chapter 2.2. Chapter 2.3 gives an overview of how we processed the raw ChIP-chip data to determine ER binding peaks. We assessed the effects of unliganded and liganded ERs at ER-binding sites, and we used the endogenous ligand E<sub>2</sub> (dual activation of ER $\alpha$  and ER $\beta$ ). In addition, to activate only ER $\alpha$  or ER $\beta$ , we used the novel, non-steroidal subtype-selective ligands PPT (ER $\alpha$  preferential activation [11]) and ERB-041 (ER $\beta$  preferential activation [12]). Figure 2.2 shows the effect of the different ligand treatments. A major aspect of our study not addressed in any of the previous ER ChIP-chip studies is the examination of the comparative binding events of ER $\alpha$  and ER $\beta$  under different ligand treatments and in breast cancer cells containing only one receptor (either ER $\alpha$  or ER $\beta$ ) or both receptors. This allowed us *(i)* to discover and map the binding sites of ER $\alpha$  and ER $\beta$  when they are present alone or together and *(ii)* to study the mechanistic roles of unliganded and liganded ER at the transcriptional level. The ER ChIP-chip study is described in Chapter 3 and published in [8].

With the popularity and reduction in ChIP-seq cost, we switched from a ChIP-chip based assay to ChIP-seq assay to examine ER $\alpha$  and ER $\beta$  chromatin binding and the recruitment of

coregulators SRC3 and RIP140 to the ERs sites (ChIP-seq methodology and enrichment analysis are described in Chapter 2.4). The benefit of ChIP-seq is that it allows us to identify ER binding sites without bias (custom-designed ChIP-chip array only allows us to examine specific, predetermined sites). At the same time, we performed several gene expression profiling studies aimed at examining ER $\alpha$ - and ER $\beta$ -mediated transcription regulation on cells expressing ER $\alpha$ -alone, ER $\beta$ -alone, and ER $\alpha$  and ER $\beta$  (gene expression profiling of ERs action is described in Chapter 2.5). The integrative study on ERs chromatin binding, coregulator recruitment, and gene expression profiling is described in Chapter 4.

## 2.2 ER CUSTOM-DESIGNED CHIP-CHIP TILING ARRAY

In the ERs ChIP-chip studies described herein, we hybridized our genomic target DNA onto custom-designed tiling arrays. The custom array was designed by employing the services of NimbleGen (<http://www.nimblegen.com/>). As transcription factor binding has been shown to be cell-line specific, care was taken to ensure that the ERs binding sites included in our arrays are found in the cell-line of interest, MCF-7 human breast cancer cells. Specifically, a literature search was done on recent published ER ChIP-chip studies, and we selected the datasets based on two criteria, (i) the cell-line used must be similar to that used in our studies, MCF-7 cells, and (ii) the ChIP-chip assays should be done in an unbiased, genome-wide manner. Due to the variation in performance between array technologies, protocols, and algorithms in calling binding sites [13], most of the current ChIP-chip studies only provide a partial picture of genome-wide binding sites. Therefore, we include, in our array, a dataset based on computational predicted ERs binding sites in MCF-7 cells. Control sites were also designed into the arrays to ensure the quality of our hybridization experiments. These control sites were designed from both

ERs non-binding chromatin regions and non-mammalian sequences. Our final ChIP-chip array tiling thus contains ~77,000 genomic regions consisting of ~61,000 ER-binding sites and ~16,000 negative/control regions. The ER-binding sites were selected based on (i) published ER ChIP-chip data [4] (10,599 sites), (ii) published ER ChIP-pair end ditag (PET) data (1234 sites) [5] and (iii) computational predicted ERE sites using an optimized algorithm [14] (37,499 sites). The breakdown of the probes regions is shown in Table 2.1. Each array consists of 38,3520 probes on them, and each probe is approximately 60 base pairs in length. On average, each binding site is covered by 5 or more probes, and the probes are tiled at a distance of ~100 base pairs from each other within the binding site, as shown in Figure 2.3. Having several probes within a binding site (the probes in each binding site is known as a probe set) increases the sensitivity of our arrays as we will be relying on the strength of the probe set signal to determine whether ER is bound to the site. Our custom-designed tiling array provides coverage of all currently known and predicted ER-binding sites across the entire genome in MCF-7 cells.

## 2.3 ER CHIP-CHIP BINDING SITES PROCESSING STEPS

Chromatin DNA fragments with an average size of 700bp were immunoprecipitated and subsequently amplified using Sigma Whole Genome Amplification Kit. 10ng of amplified DNA was used for each ChIP-chip assay. The ER ChIP-chip assays were run as two-color experiments with the experimental (ChIP) DNA sample labeled with dyes of a certain color (Cy5) and the control (background) DNA sample labeled with another dye of different color (Cy3) according to NimbleGen instructions. The two samples are then co-hybridized onto the array using a MAUI hybridization unit and scanned. The signal intensities of the samples (experimental and control) at each probe are extracted from the scanned images of the arrays. Enrichment at the probe is

quantified by probe intensity  $\log_2$ -ratios ( $P_j$ ), calculated as the base 2 logarithm of the ratio between the experimental and control samples.

$$P_j = \log_2 \frac{T_j}{R_j}$$

The probe  $\log_2$ -ratio values ( $P_j$ ) are then scaled to center the ratio data around zero. Scaling is performed by subtracting the Turkey bi-weight mean for the  $\log_2$ -ratio values for all probes on the array from each probe's  $\log_2$ -ratio value. Turkey bi-weight mean is used to identify the center of the data, as it is less sensitive to outliers, i.e., outliers far from the median contribute little to the average. Therefore, each probe on the array will have a corresponding scaled  $\log_2$ -ratio,  $S_j$  value.

$$\gamma_j = \log_2 \frac{T_j}{R_j} - S_p$$

$j = 1, 2, \dots, 383520$   
 $T_j =$  test sample  
 $R_j =$  reference sample

where turkey bi-weight value,  $S_p$ ,

$$S_p = \frac{\sum_{j=1}^{383520} y_j (1 - y_j^2)^2}{\sum_{j=1}^{383520} (1 - y_j^2)^2}$$

$M = \text{median}(p_1, \dots, p_{383520}),$   
 $MAD(p) = \text{median}(|p_1 - M|, \dots, |p_{383520} - M|),$   
 $y_j = \frac{p_j - M}{S * MAD(p)}$

Transcription factors binding sites are characterized by their localized response in the genome and by their enrichment in the experimental sample compared to the control sample (consecutive probes with positive scaled  $\log_2$ -ratio values in a typical ChIP-chip experiment). Therefore, peaks (potential binding sites) are subsequently detected by searching for four or more probes whose scaled  $\log_2$ -ratio values,  $S_j$ , are above a specified cutoff value within a 700 base-pair sliding window region (see Figure 2.4). We then use a permutation-based approach to find statistically significant peaks that are likely to be representative of actual ERs binding events. This analysis estimates the false discovery rate (FDR) for each peak. The FDR value is equal to the probability of finding a peak of similar significance by chance and is estimated by repeatedly

(20 times) randomly permuting the scaled  $\log_2$ -ratio values and searching for peaks. In general, for a ChIP-chip experiment, the lower the FDR score, the more likely the peak corresponds to a transcription factor binding site (TFBS).

In order to determine a FDR cutoff value suitable for our ChIP-chip studies, we conducted a negative control ChIP-chip experiment in which we immunoprecipitated ER $\beta$  in normal MCF-7 breast cancer cells and subsequently hybridized the samples on the array. Normal MCF-7 cells only contain ER $\alpha$  protein, and therefore there should not be any binding sites in this ChIP-chip experiment. By using a FDR cutoff of 0.2, only 17 peaks were called as binding sites out of a potential  $\sim 77000$  binding regions (see Figure 2.5). This is much lower than the accepted error rates in current whole genome ChIP-chip experiment. Hence, for our ChIP-chip studies, we used a FDR cutoff of less than 0.2 as indicative of ER binding site.

#### 2.4 CHIP-SEQUENCING OF ER $\alpha$ , ER $\beta$ , SRC3, AND RIP140 BINDING SITES

The first next-generation sequencing platform was introduced to the research community in 2006 by the 454 Corporation. The machine could generate around 200 thousands sequencing reads at a read length of approximately 100 base pairs (bp). Suddenly, sequencing became much cheaper and faster compared to what was possible with Sanger sequencing. Towards the end of 2006, Illumina introduced the Genome Analyzer (GA) sequencing machine. The sequencing capacity of the GA was impressive as it was able to read up to 80 million DNA template clusters simultaneously at an average length of 36-50bp [15, 16]. Although the Illumina sequencing machine has a much shorter read length than the 454 instrument, it is extremely well suited for sequencing ChIP samples. This is because ChIP DNA fragments are short sequences; so, when it

is coupled with Illumina massive parallel reading capacity, ChIP DNA libraries can be sequenced and the required data generated at an extreme fast rate.

Our initial study of ER $\alpha$  and ER $\beta$  binding sites was done using a custom-designed ChIP-chip tiling array. The probes on the tiling array were designed based on known ER $\alpha$  sites, as well as other presumed ER binding sites predicted computationally (Chapter 2.2). Thus, the ChIP-chip approach was limited to some degree in its use to study ER $\beta$  binding, because some ER $\beta$  unique binding sites might have been missed. In our second study (Chapter 4), we used the increased speed of the ChIP-seq method to study, in a genome-wide unbiased manner, the chromatin binding of both ER $\alpha$  and ER $\beta$ , as well as the recruitment of the coregulators SRC3 and RIP140 to the ER sites. All our ChIP-seq libraries were sequenced on the Illumina platform.

The ChIP-seq libraries were prepared following the manufacturer's protocols. Briefly, chromatin DNA fragments with an average size of 700bp were immunoprecipitated. The ends of the DNA fragments were converted into phosphorylated blunt ends by T4 DNA polymerase, Klenow enzyme, and T4 polynucleotide kinase. Illumina adaptors were subsequently ligated to the ChIP DNA fragments. The DNA fragments were subjected to 15 cycles of PCR amplification. The fraction of fragments averaging 200bp was selectively cut out from a 2% agarose gel and eluted using a Qiagen gel extraction kit. Using the Illumina platform, 36bp length tags were sequenced. The raw sequencing image data were processed by the Illumina analysis pipeline and mapped uniquely onto the human reference genome (NCBI v36, hg18) using ELAND (Efficient Large-scale Alignment of Nucleotide Databases).

## 2.5 PEAK CALLING FOR CHIP-SEQUENCING TAGS

After mapping the raw sequenced tags to the genome by ELAND, we needed to identify the chromatin binding sites. Binding sites are found by looking for ‘peaks’ – regions of significant tag enrichment. There are now a large number of peak-calling programs, both public and commercial, that were developed for ChIP-seq [17]. We selected CCAT [18], which was developed at Genome Institute of Singapore for peak calling. In addition to sequencing the ChIP library for our proteins of interest (ER $\alpha$ , ER $\beta$ , SRC3, and RIP140), we also sequenced negative control libraries. Negative control libraries, which are also known as Input libraries, are generated by sequencing ‘input DNA’ (non-ChIP genomic DNA). As several batches of cells were thawed and used to generate the ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 ChIP libraries, an Input control library was also generated for each batch of cells. This is done to control for cell batch variations. CCAT requires both the ChIP library and Input library for peak calling. Based on the comparison of the ChIP library and the Input library, CCAT will estimate the confidence of the peaks called (quantified by False Discovery Rate (FDR)). For comparison, besides using the corresponding Input library to estimate the confidence of the peaks called by CCAT for each ChIP library, we also used the other Input libraries. To our surprise, we found that the number of peaks called for each ChIP library differed significantly (~6000 to 90000 binding sites between different Input libraries) when using different Input libraries. For example, Table 2.2 shows the number of peaks called by CCAT for ER $\beta$  [ $\beta$  cells] ChIP library across different Input libraries (Input\_2 is the corresponding Input library generated from the same batch of cells as ER $\beta$  [ $\beta$  cells] ChIP). The differences in number of peaks called might be caused by cell batch variations since the Input libraries were generated from different batch of cells or caused by CCAT peak calling algorithm. Upon closer inspection of CCAT results, we noticed that CCAT was using



discrete FDR values. For example, the FDR value for peaks in ER $\beta$  [ $\beta$  cells] library jumped from FDR of 0.055 to FDR 0.09 (see Table 2.3) when the FDR was sorted in ascending order. Hence, by using a certain FDR as cutoff, the number of peaks called for a particular ChIP library can be different when using different Input libraries, as the FDR values are discrete and not continuous.

To further confirm that the problem is specifically for CCAT, we used another peak calling program, MACS (Model-based Analysis of ChIP-sequencing) [19]. MACS is one of the most popular peak calling programs currently used by the ChIP-seq community. As shown in Table 2.4 (using default MACS cutoff), the number of peaks for ER $\beta$  [ $\beta$  cells] is very similar across the different Input libraries. This confirms that the variation in the number of peaks across different Input libraries (Table 2.2) is caused by the CCAT algorithm which uses discrete FDR values. Hence, identification of enriched sites in our sequenced ChIP tags (Chapter 4) was performed using MACS software by comparing each ChIP library to the corresponding Input library as a control. We defined binding peaks as those above the p-value cutoff of  $6e-7$  and FDR rate of 0.01.

## 2.6 TRANSCRIPTIONAL PROFILING OF ER $\alpha$ AND ER $\beta$ ACTIONS

As the ER status in breast tumors is an important prognostic factor of response to endocrine therapies, there is much interest in understanding ER-targeted genes. The introduction of DNA microarray technology has enable researchers to examine the expression of thousands of genes. Both Affymetrix GeneChip and Illumina BeadChip are popular choices because of the large number of probes on the arrays and the well-annotated sequences. In addition, many tools have been developed for the analysis of the data generated from GeneChip and BeadChip.

We performed several gene expression profiling studies, using Affymetrix GeneChip U133 plus 2.0, aimed at dissecting ER $\alpha$ - and ER $\beta$ -mediated transcription regulation in cells expressing ER $\alpha$ -alone, ER $\beta$ -alone, and ER $\alpha$  and ER $\beta$  together. Total RNA was used to generate cRNA, which was labeled with biotin according to techniques recommended by Affymetrix. The biotin-labeled cRNA was then hybridized to Affymetrix U133 plus 2.0 GeneChips, which contain oligonucleotide probe sets for over 47000 transcripts. After washing, the chips were scanned and analyzed using Affymetrix processing software. CEL files were processed using GeneSpring GX 11.0 software (Agilent) to obtain fold-change and p-value (with Benjamini and Hochberg multiple test correction) for each gene and for each treatment relative to the vehicle control. We considered genes with fold-change > 1.5 and p-value < 0.05 as differential expressed.

## 2.7 TABLES AND FIGURES

**Table 2.1.** Binding sites represented in custom-designed arrays.

	Validated ER binding sites		GIS Dtags data				Prediction	Brown's data	Controls			
	Binding	Non-binding	moPET1-2	moPET3-5	moPET6+	Prokaryotes			Exons	PET1	Neg regions	
Binding sites	86	68	6100	1171	303	37499	10599	100	300	10661	5109	
Total Binding Regions	154		7574				37499	10599	16170			
Total Probes	76998						183458	60894	62170			

**Table 2.2.** Number of peaks called by CCAT for ER $\beta$  [ $\beta$  cells] library.

<b>Input Library</b>	<b>No of peaks</b>
Input_1	240709
Input_2	150955
Input_3	190336
Input_4	210495

Four Input libraries from different batches of cells were used as control by CCAT. (Input\_2 is generated from the same batch of cells used for ER $\beta$  [ $\beta$  cells] ChIP)

**Table 2.3.** FDR values of the peaks called by CCAT for ER $\beta$  [ $\beta$  cells] library (control is Input\_2).

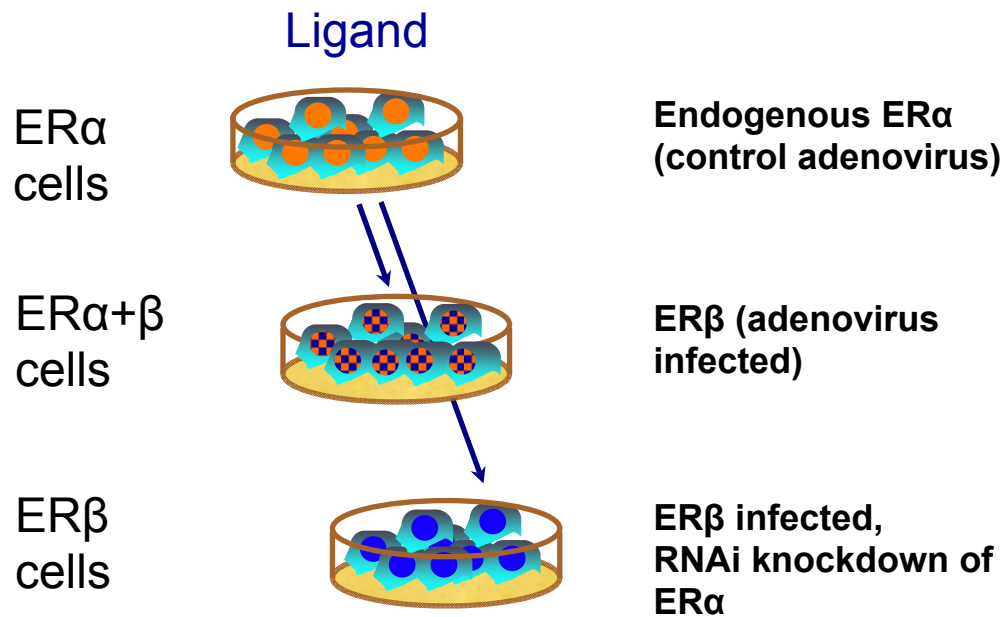
Chromosome	Peak location	local FDR (ascending order)
chr10	109546445	0.055
chr20	34647485	0.055
chr10	119866655	0.055
chr13	35186545	0.055
chr5	29416435	0.055
chr7	1928775	0.055
<b>chr13</b>	<b>112151625</b>	<b>0.055</b>
<b>chr2</b>	<b>132743255</b>	<b>0.09</b>
chr3	63833095	0.09
chr20	48720555	0.09
chr5	127256665	0.09
chr2	212093375	0.09



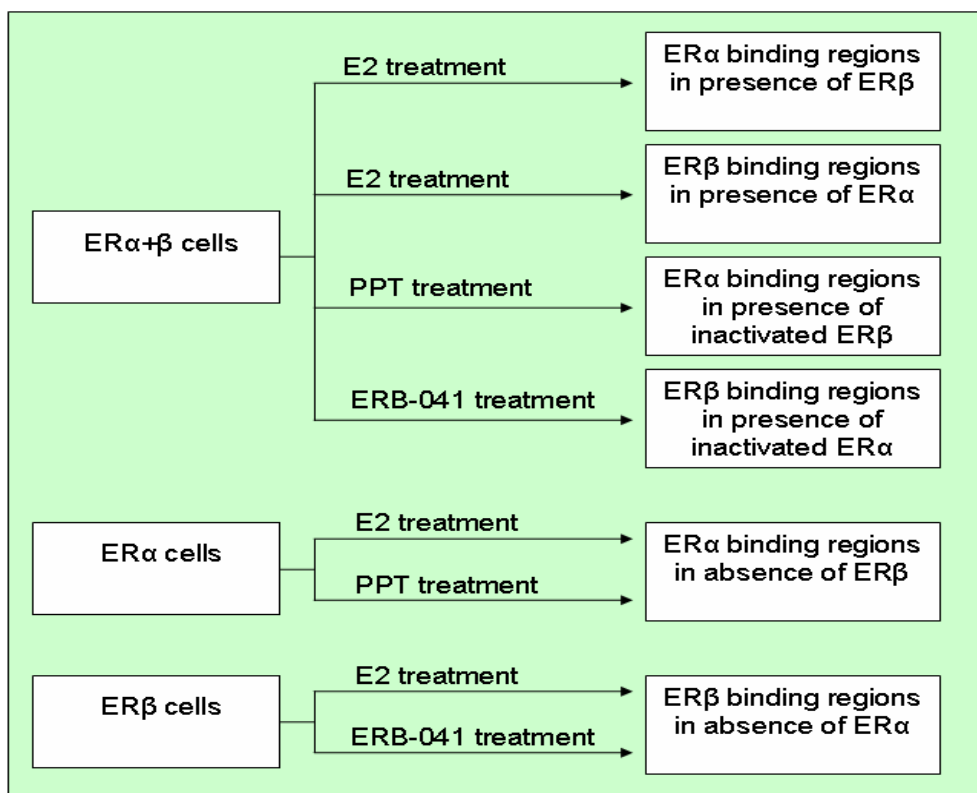
**Table 2.4.** Number of peaks called by MACS for ER $\beta$  [ $\beta$  cells] library.

<b>Input Library</b>	<b>No of peaks</b>
Input_1	84300
Input_2	75547
Input_3	86658
Input_4	82705

Four Input libraries from different batches of cells were used as control by MACS. (Input\_2 is generated from the same batch of cells used for ER $\beta$  [ $\beta$  cells] ChIP)

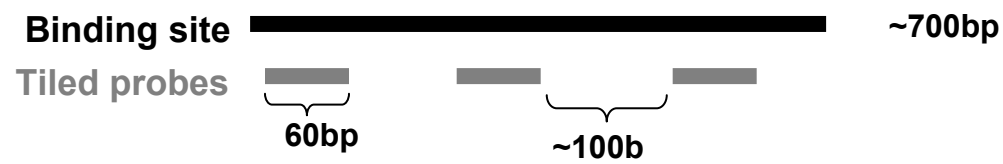


**Figure 2.1.** MCF-7 cells were infected with control  $\beta$ -galactosidase adenovirus or ER $\beta$ -containing adenovirus to generate cells containing ER $\alpha$ -only and ER $\alpha$ +ER $\beta$ , respectively. Cells containing ER $\beta$ -only were generated by the subsequent knockdown of ER $\alpha$  by siRNA transfection of cells containing ER $\alpha$ +ER $\beta$ .

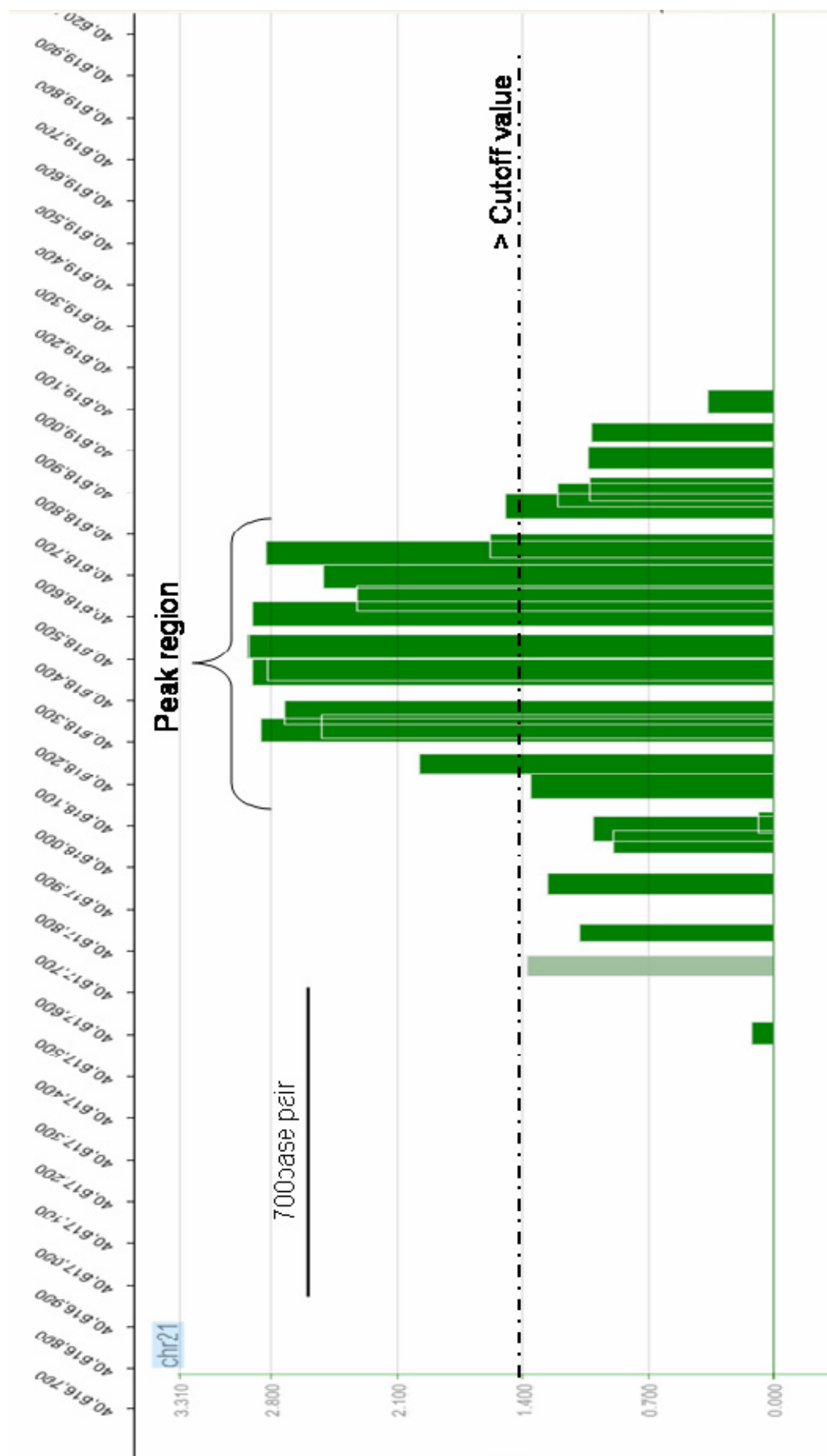


**Figure 2.2.** The effects of different ligands on either ERα or ERβ binding sites were assessed. The ligands used in our studies are endogenous ligand E<sub>2</sub> (dual activation of ERα and ERβ) and the novel, non-steroidal ligands PPT (ERα preferential activation) and ERB-041 (ERβ preferential activation).

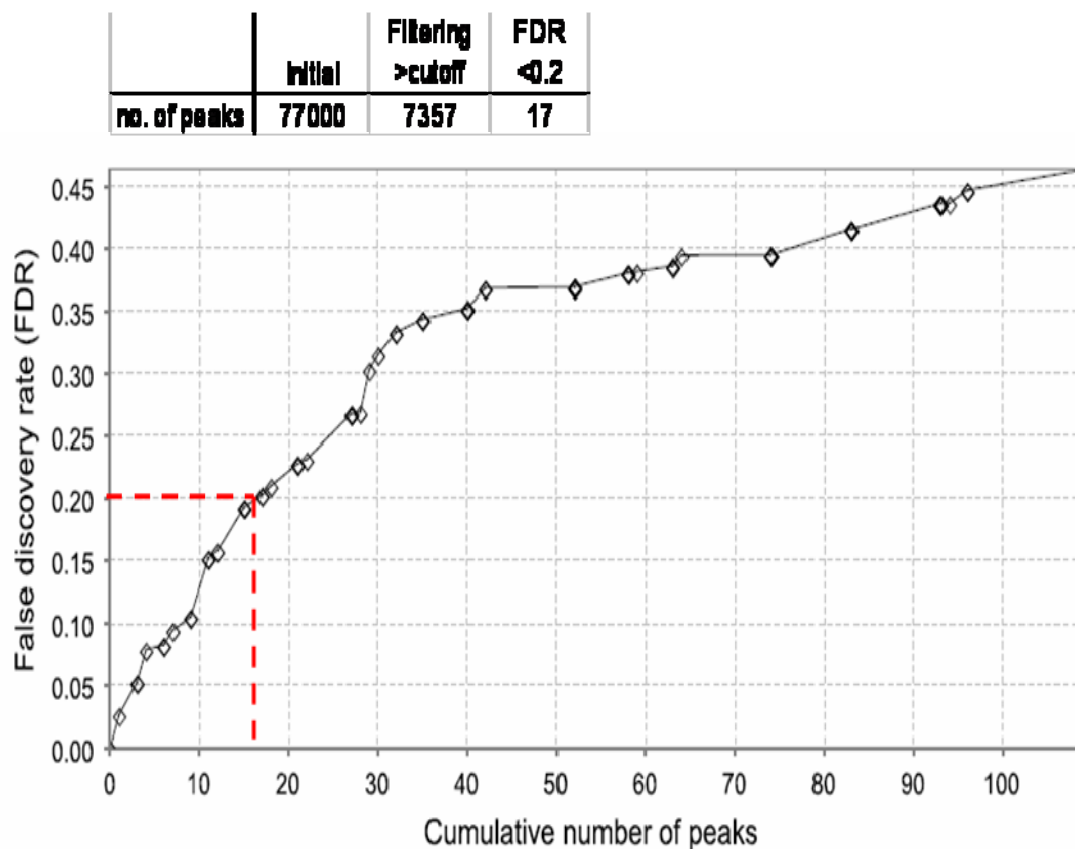




**Figure 2.3.** Schematic diagram showing locations of tiled probes in the custom-designed tiling array. Each probe is 60 bp in length, and probes are tiled ~100 bp from each other.



**Figure 2.4.** Genomic regions are identified as peaks (potential binding sites) if there are four or more probes whose scaled log<sub>2</sub>-ratio values are above a specified cutoff value within a 700 base pairs sliding window region.



**Figure 2.5.** A negative control ChIP-chip experiment was done to determine the most suitable FDR cutoff value for use in our studies. MCF-7 cells were treated with  $E_2$  and immunoprecipitated against ER $\beta$ . By using a FDR cutoff of 0.2, we detected 17 “potential” binding sites (out of the theoretical 0 binding sites in this experiment), which is much lower than the current accepted error rates for a ChIP-chip experiment.

## 2.8 REFERENCES

1. Bajic, V.B., et al., *Dragon ERE Finder version 2: A tool for accurate detection and analysis of estrogen response elements in vertebrate genomes*. Nucleic Acids Res, 2003. **31**(13): p. 3605-7.
2. Lin, C.Y., et al., *Discovery of estrogen receptor alpha target genes and response elements in breast tumor cells*. Genome Biol, 2004. **5**(9): p. R66.
3. Tang, S., et al., *Computational method for discovery of estrogen responsive genes*. Nucleic Acids Res, 2004. **32**(21): p. 6212-7.
4. Carroll, J.S., et al., *Genome-wide analysis of estrogen receptor binding sites*. Nat Genet, 2006. **38**(11): p. 1289-97.
5. Lin, C.Y., et al., *Whole-genome cartography of estrogen receptor alpha binding sites*. PLoS Genet, 2007. **3**(6): p. e87.
6. Liu, Y., et al., *The genome landscape of ERalpha- and ERbeta-binding DNA regions*. Proc Natl Acad Sci U S A, 2008. **105**(7): p. 2604-9.
7. Kininis, M., et al., *Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters*. Mol Cell Biol, 2007. **27**(14): p. 5090-104.
8. Charn, T.H., et al., *Genome-wide dynamics of chromatin binding of estrogen receptors alpha and beta: mutual restriction and competitive site selection*. Mol Endocrinol. **24**(1): p. 47-59.
9. Zhao, C., et al., *Genome-wide mapping of estrogen receptor-beta-binding regions reveals extensive cross-talk with transcription factor activator protein-1*. Cancer Res. **70**(12): p. 5174-83.
10. Chang, E.C., et al., *Estrogen Receptors alpha and beta as determinants of gene expression: influence of ligand, dose, and chromatin binding*. Mol Endocrinol, 2008. **22**(5): p. 1032-43.
11. Stauffer, S.R., et al., *Pyrazole ligands: structure-affinity/activity relationships and estrogen receptor-alpha-selective agonists*. J Med Chem, 2000. **43**(26): p. 4934-47.
12. Malamas, M.S., et al., *Design and synthesis of aryl diphenolic azoles as potent and selective estrogen receptor-beta ligands*. J Med Chem, 2004. **47**(21): p. 5021-40.
13. Johnson, D.S., et al., *Systematic evaluation of variability in ChIP-chip experiments using predefined DNA targets*. Genome Res, 2008. **18**(3): p. 393-403.
14. Vega, V.B., et al., *Multiplatform genome-wide identification and modeling of functional human estrogen receptor binding sites*. Genome Biol, 2006. **7**(9): p. R82.
15. Barski, A., et al., *High-resolution profiling of histone methylations in the human genome*. Cell, 2007. **129**(4): p. 823-37.
16. Johnson, D.S., et al., *Genome-wide mapping of in vivo protein-DNA interactions*. Science, 2007. **316**(5830): p. 1497-502.
17. Liu, E.T., S. Pott, and M. Huss, *Q&A: ChIP-seq technologies and the study of gene regulation*. BMC Biol. **8**: p. 56.
18. Xu, H., et al., *A signal-noise model for significance analysis of ChIP-seq with negative control*. Bioinformatics. **26**(9): p. 1199-204.
19. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS)*. Genome Biol, 2008. **9**(9): p. R137.

### **CHAPTER 3**

#### **LIGAND REGULATION OF CHROMATIN BINDING SITES OF ESTROGEN RECEPTORS $\alpha$ AND $\beta$ : SUBTYPE COMPETITION, MUTUAL RESTRICTION, AND LIGAND SELECTIVITY**

##### **3.1 ABSTRACT**

Estrogens exert profound effects on the gene expression and biological response programs of their target cells, including breast cancer cells. Because these effects on gene expression are mediated by two nuclear hormone receptors, estrogen receptor  $\alpha$  (ER $\alpha$ ) and estrogen receptor  $\beta$  (ER $\beta$ ), that can coexist in the same cells, we explore, in this report, the chromatin binding sites for ER $\alpha$  and ER $\beta$ , and the interplay of these two receptors when they are present separately or together in breast cancer cells. For these studies, we use MCF-7 human breast cancer cells with three different complements of these receptors (ER $\alpha$  only, ER $\beta$  only, and ER $\alpha$  and ER $\beta$ ), treatment of cells with estradiol (E<sub>2</sub>), which binds well to both ERs, or with subtype-specific ligands that bind to only one of the ERs, and chromatin immunoprecipitation (ChIP)-chip analysis with a custom-designed tiling array for ER binding sites across the genome to examine the effects of ligand-occupied and unoccupied ER $\alpha$  and ER $\beta$  on chromatin binding as determinants of gene expression. Our studies demonstrate that there is substantial overlap in the chromatin binding sites for E<sub>2</sub>-liganded ER $\alpha$  and ER $\beta$  when they are present alone in cells, but that many fewer sites are shared when both ERs are present together. Although each ER subtype restricts the binding site occupancy of the other subtype, overall, ER $\alpha$  appears to dominate the binding of ER $\beta$ , although this is observed only at stimulated genes but not at repressed genes. The binding site regions of both ER $\alpha$  and ER $\beta$  are markedly enriched in estrogen response element (ERE) sequence motifs, and when both ERs are present together, ER $\alpha$  appears to displace ER $\beta$  so that ER $\beta$  binds to sites substantially less enriched in EREs. Studies with ER subtype-specific ligands reveal that it is the liganded ER subtype that principally determines the

sites of chromatin binding. These findings highlight the dynamic interplay between the two ER subtypes in their selection of chromatin binding sites, how this is modulated by their state of ligand occupancy, and their impact on gene expression.

### 3.2 INTRODUCTION

Estrogens play key roles in many aspects of reproductive physiology, development, and metabolism, and they are also involved in several disease states, including breast and endometrial cancers [1, 2]. The effects of estrogens are mediated through two estrogen receptors, estrogen receptor alpha (ER $\alpha$ ) and beta (ER $\beta$ ), that function as ligand-modulated transcription factors, up- and down-regulating gene expression in a target tissue-selective manner [3, 4]. The presence of ER $\alpha$  in breast cancer cells and in various tissues is associated with enhanced proliferation in response to estrogens, whereas several studies have implicated ER $\beta$  as exerting antiproliferative effects [5-9].

ER $\alpha$  and ER $\beta$  are highly homologous in their DNA-binding domains (97% identity), but they are quite different in their ligand-binding domains (56% identity) and transcriptional activation function-1 (18% identity) domains. The differences in their ligand-binding domains allow the two ER subtypes to bind certain ligands with high selectivity, for one or the other ER subtype [10-13]. Although most human breast cancers co-express both ERs [14-16], much less is known about the role of ER $\beta$  in breast cancer and how the presence of both ERs might affect cellular responses to estrogen, although the presence of ER $\beta$  in breast tumors is generally associated with a better prognosis [16-20].

Recently, the chromatin binding sites for the ER $\alpha$  have been mapped on a genome-wide scale using chromatin immunoprecipitation (ChIP) combined with either DNA microarray (ChIP-chip)

or sequencing (ChIP-PET) analysis [21-24]. These studies have provided an unprecedented view of the diversity and distribution of ER binding sites throughout the genome. Unfortunately, most of these studies examined only ER $\alpha$  binding, with the exception of the study done by Liu et al. [24], that investigated binding sites in only a portion of the genome (chromosomes 1, 3, 6, 21, 22, X, and, Y) in cells treated with the ER subtype non-selective ligand, E<sub>2</sub>. Clearly, additional studies are needed to examine the characteristics of ER $\beta$  binding sites throughout the whole genome and to investigate whether ER $\alpha$  and ER $\beta$  collaborate and/or compete for these binding sites in the presence of E<sub>2</sub> and various ER subtype-specific ligands.

ERs normally function as dimers at chromatin binding sites where they recruit a variety of coregulators, histone-modifying enzymes, and other factors to up or downregulate the transcription of hundreds of genes that markedly influence cell phenotype [3-5, 25]. It is now well documented that ER $\alpha$  and ER $\beta$  can exist in cells as homodimers when present alone, and additionally as heterodimers when present together [26, 27]. Therefore, to better understand the interplay between ER $\alpha$  and ER $\beta$  binding at the genomic level, we have used ChIP-chip analysis with a custom-designed tiling array to investigate the binding of ER $\alpha$  and ER $\beta$ , when present together or separately, at all documented high probability and hypothetical estrogen receptor binding sites across the genome in MCF-7 breast cancer cells. Specifically, in this work, we have used subtype-selective vs. nonselective ligands to examine the effects of ligand-occupied and unoccupied ER $\alpha$  and ER $\beta$  on chromatin binding events and as determinants of gene expression in breast cancer cells. To this end, we have profiled the genome-wide binding events of the two ERs in cells containing various complements of ER $\alpha$  and ER $\beta$ , treated with either 17 $\beta$ -estradiol (E<sub>2</sub>) or with the ER $\alpha$ -selective agonist PPT [13] or the ER $\beta$ -selective agonist ERB-041 [11]. Our ChIP-chip studies extend beyond previous large-scale ER binding experiments [21, 22, 24] and

demonstrate that there is substantial overlap in the chromatin binding sites for E<sub>2</sub>-liganded ER $\alpha$  and ER $\beta$  when they are present alone in cells, but that many fewer sites are shared when both ERs are present together. Although each ER subtype restricts the binding site occupancy of the other ER, overall, ER $\alpha$  appears to dominate the binding of ER $\beta$ . We find that the binding site regions of both ER $\alpha$  and ER $\beta$  are markedly enriched in estrogen response element (ERE) sequence motifs, but when both ERs are present together, ER $\alpha$  appears to displace ER $\beta$ .

### 3.3 MATERIALS AND METHODS

#### *Ligands, Cell Culture and Adenovirus Infection*

MCF-7 cells were cultured in MEM (Sigma, St Louis, MO), supplemented with 5% calf serum (HyClone, Logan, UT), and 100  $\mu$ g/ml penicillin/streptomycin (Invitrogen, Carlsbad, CA). For estrogen-free experiments, the cells were maintained in phenol red-free MEM plus 5% charcoal-dextran-treated calf serum for at least 3 days and were then seeded at a density of  $3 \times 10^5$  cells per 10 cm tissue culture dish (Corning, Corning, NY) for 2 days before adenovirus infection. Recombinant adenoviruses were constructed and prepared as described [3]. Cells were infected with either control adenovirus expressing  $\beta$ -galactosidase (Ad) or adenovirus containing ER $\beta$  (AdER $\beta$ ) for 72h before ligand treatment. Conditions used were those described previously [3, 9] to generate MCF-7 cells expressing levels of ER $\beta$  equal to that of the endogenously expressed ER $\alpha$ . Estradiol was from Sigma. The ER subtype-selective ligands, PPT and ERB-041, were synthesized as described [13, 28]. Studies used 10 nM E<sub>2</sub>, 50 nM PPT, and 500 nM ERB-041, concentrations that reflect their relative binding affinities, and give maximal occupancy of receptors by these ligands.

#### *siRNA Transfection*



siRNA experiments were performed as previously described, and resulted in knockdown of ER $\alpha$  mRNA and protein by greater than 95% [4]. siER $\alpha$  sequences (Dharmacon) were: forward, 5'-UCAUCGCAUCCUUGCAAAdTdT-3', and reverse, 5'-UUUGCAAGGAAUGCGAUGAdTdT-3'.

### *ChIP Assays*

ChIP for ER $\alpha$  and ER $\beta$  were carried out as described [29] and used the following antibodies: ER $\alpha$  antibody HC-20 (Santa Cruz Biotechnology); ER $\beta$  antibodies were a combination with equal parts of CWK-F12 (produced by our lab) [30], GTX70182 (GeneTex), GR40 (Calbiochem), and PA1-311 (Affinity Bioreagents). The ChIP DNA was used for ChIP-chip analysis and quantitative real-time PCR.

### *ChIP-chip Analyses*

The ChIP DNA was hybridized on a custom-designed tiling array produced by NimbleGen. The design of the tiling array is described in details in Chapter 2.2.

## 3.4 RESULTS

### 3.4.1 Use of Custom Tiled Microarray of ER Binding Sites

The custom-designed tiling arrays used in our studies (design methodology described in Chapter 2.2) were designed to provide coverage of all known and predicted ER-binding sites across the entire genome in MCF-7 cells. MCF-7 cells having the three complements of ER (ER $\alpha$  only, ER $\alpha$ +ER $\beta$ , ER $\beta$  only) were grown in the absence of hormones for at least three days and were then treated with control, 0.1% ethanol vehicle or one of the three ligands for a short period of time (45 min). Chromatin fragments bound by ER were immunoprecipitated and hybridized onto our tiling arrays. We performed three biological replicates (each biological

replicate from an independent experiment consisted of two separate hybridizations, technical replicates, onto the tiling arrays) to identify enriched binding sites. The raw intensity signals of the ChIP-chip experiments were normalized and averaged across the three replicates. The binding sites were then identified using the algorithms described in Chapter 2.3. Briefly, ERs binding sites are identified from the ChIP-chip results by the intersection of peaks detection (4 or more probes whose intensity signals are above a specific threshold) and default false discovery rate (FDR) cutoff. The raw ChIP-chip data were processed, and the peak cutoff threshold and FDR values were calculated using NimbleGen recommended software, NimbleScan.

Table 3.1 shows the number of ER $\alpha$  and ER $\beta$  binding sites under the various experimental conditions. In cells containing ER $\alpha$  only, we probed by ChIP for ER $\alpha$  binding sites (hereafter designated ER $\alpha$ [ $\alpha$  cells] binding sites); in cells with ER $\beta$  only, we probed for ER $\beta$  binding sites (ER $\beta$ [ $\beta$ -cells] binding sites), and in cells containing both ER $\alpha$  and ER $\beta$ , we probed for both ER $\alpha$  binding sites (ER $\alpha$ [ $\alpha\beta$  cells] binding sites) and for ER $\beta$  binding sites (ER $\beta$ [ $\alpha\beta$  cells] binding sites). A number of mock ChIP-chip experiments were also performed to ensure the fidelity of our ChIP-chip analyses and to determine the FDR score cutoff (Chapter 2.3). With E<sub>2</sub> treatment, in MCF-7 cells expressing ER $\alpha$  only, we identified 4405 ER $\alpha$ -binding sites (ER $\alpha$ [ $\alpha$  cells] sites); in MCF-7 expressing ER $\beta$  only, we identified 1897 ER $\beta$ -binding sites (ER $\beta$ [ $\beta$  cells] sites), and in MCF-7 cells expressing both ER $\alpha$  and ER $\beta$ , we identified 3252 ER $\alpha$ -binding sites (ER $\alpha$ [ $\alpha\beta$  cells] sites) and 1744 ER $\beta$ -binding sites (ER $\beta$ [ $\alpha\beta$  cells] sites) (See Table 3.1).

### 3.4.2 Validation of ER $\alpha$ and ER $\beta$ ChIP-Chip Results

To confirm the validity of our ChIP-chip dataset, we examined recruitment of the ERs, in the three cell-types, on a set of randomly selected binding sites from our ChIP-chip dataset using

ChIP-qPCR with antibodies against ER $\alpha$  or ER $\beta$  (Figure 3.1). We tested a total of 42 sites and validated ER binding in ~91% (39 of 42) of the selected sites. In this respect, our ChIP-chip experiments had false-positive error rates of approximately 10%, which are similar to the error rates reported in other genome-wide ChIP-chip studies [21, 22].

### 3.4.3 ER $\alpha$ has a Dominant Effect on the Distribution of ER Binding

To understand how the presence of the ER-subtype partner might influence the pattern of chromatin binding sites for ER $\alpha$  [in  $\alpha$  cells] and ER $\beta$  [in  $\beta$  cells], we compared the pattern of binding site occupancy by these ERs after E<sub>2</sub> exposure in the three cell types. Specifically, we compared how ER $\alpha$  binding sites in ER $\alpha$  cells changed when ER $\beta$  was also present (i.e., ER $\alpha$ [ $\alpha$  cells] sites vs. ER $\alpha$ [ $\alpha\beta$  cells] sites); and conversely, we examined how ER $\beta$  binding sites in ER $\beta$  cells changed when ER $\alpha$  was also present (ER $\beta$ [ $\beta$ -cells] sites vs. ER $\beta$ [ $\alpha\beta$  cells] sites). The results are visualized in the Venn diagram in Figure 3.2, showing the shift in binding sites occupancy with the addition of the other ER subtype partner.

Intriguingly, in cells containing both ER $\alpha$  and ER $\beta$  treated with E<sub>2</sub>, the presence of ER $\beta$  restricts the range of binding sites for ER $\alpha$  when compared to that seen in ER $\alpha$ -only cells (3252 ER $\alpha$ [ $\alpha\beta$  cells] sites vs. 4405 ER $\alpha$ [ $\alpha$  cells] sites). Despite the decrease in the number of ER $\alpha$ [ $\alpha\beta$  cells] binding sites, the overlap between ER $\alpha$ [ $\alpha\beta$  cells] and ER $\alpha$ [ $\alpha$  cells] binding sites is very high, with 92% of ER $\alpha$ [ $\alpha\beta$  cells] binding sites overlapping the ER $\alpha$ [ $\alpha$  cells] binding sites (Figure 3.2A). By contrast, as shown in Figure 3.2B, ER $\beta$ [ $\alpha\beta$  cells] binding sites overlap with only a small number of ER $\beta$ [ $\beta$  cells] binding sites (806, ~40%). It appears that, in cells containing both ERs the presence of E<sub>2</sub>-liganded ER $\beta$  has much less influence on ER $\alpha$  binding site selection than does E<sub>2</sub>-liganded ER $\alpha$  have on the binding site profile of ER $\beta$ ; in fact, the presence of ER $\alpha$

prevented ER $\beta$  from accessing the majority of the binding sites it accessed in ER $\beta$ -only cells. These data suggest that ER $\alpha$  is more “dominant” in competing for ER binding sites than ER $\beta$  when the two ER subtypes are coexpressed in the same cells. Curiously, in the presence of ER $\alpha$ , ER $\beta$  bound to 938 new sites, in fact, to more new sites than to sites it accessed in ER $\beta$ -only cells. While addition of ER $\beta$  caused ER $\alpha$  to bind to some new sites, these represented less than 10% of the ER $\alpha$ [ $\alpha$  cell] sites.

#### 3.4.4 Mutual Competition Between ER $\alpha$ and ER $\beta$ Binding Site Occupancy Restricts the Number of Potential Heterodimer Sites

To further evaluate the influence of each ER subtype partner on chromatin binding, we examined the occupancy of ER binding sites by ER $\alpha$  and ER $\beta$  when they are present either separately or together in cells (Figure 3.3A and 3.3B, respectively) after E<sub>2</sub> treatment. When present separately in cells, ER $\alpha$  or ER $\beta$  can each occupy many of the same sites: About 73% of ER $\beta$ [ $\beta$  cells] binding sites correspond to ER $\alpha$ [ $\alpha$  cells] binding sites, though the converse is less, with ~31% of the ER $\alpha$ [ $\alpha$  cells] binding sites also being accessible to ER $\beta$  in ER $\beta$ -cells. This is consistent with the current knowledge that ER $\alpha$  and ER $\beta$  can recognize the same estrogen response element motif, and is also consistent with our current observation that ER $\alpha$  appears to be the dominant binding subtype. Notably, 28% of the combined total of ER $\alpha$ [ $\alpha$  cells] sites plus ER $\beta$ [ $\beta$  cells] sites represent sites that are *in common* to both ERs (i.e., can be occupied by *either* ER $\alpha$  or ER $\beta$  when present alone). Because both ER $\alpha$  and ER $\beta$  have been shown to interact with each other and to bind to DNA as heterodimers [31, 32], the ER binding sites common to both ER $\alpha$  and ER $\beta$  (Figure 3.3A) may represent *potential* sites for ER $\alpha$ /ER $\beta$  heterodimer binding.

We next investigated the association of ER $\alpha$  and ER $\beta$  binding sites when both receptors are present in the cells [ $\alpha\beta$  cells]. As shown in Figure 3.3B, the ER $\alpha$  [ $\alpha\beta$  cells] sites and the ER $\beta$  [ $\alpha\beta$  cells] sites are much more distinct than were the ER $\alpha$  [ $\alpha$  cells] sites and ER $\beta$  [ $\beta$  cells] sites. Thus, the number of binding sites that can be shared when both ER $\alpha$  and ER $\beta$  are present (579) is much less than the number of sites that are common (1386) (i.e., can be occupied by either ER $\alpha$  or ER $\beta$  when present alone). The *shared binding sites* represent only 13% of the total sites in the  $\alpha\beta$  cells, whereas the *binding sites in common* to  $\alpha$  cells and  $\beta$  cells represented 28% of the total binding sites.

The large reduction in the number of ER $\alpha$  and ER $\beta$  binding sites that are shared when both ER subtypes are present together suggests again that there is active competition between the two receptors for binding at many chromatin sites: When both receptors are present separately, they can bind to many of the same sites, because there is no competition from the other ER subtype. When they are present together in cells, however, ER $\alpha$  and ER $\beta$  compete with each other for binding to these common sites, and thus end up sharing a more limited number of binding sites. Thus, the number of *likely* ER $\alpha$ -ER $\beta$  heterodimerization sites, measured when both ER subtypes are present (579 shared sites), is much smaller than the number of *potential* ER $\alpha$ -ER $\beta$  heterodimerization sites (1386 common sites), measured by the overlap of sites for each ER alone. Overall, these results clearly implicate a dynamic binding paradigm for ER $\alpha$  and ER $\beta$  in the presence of their subtype partner and highlight that the two ERs can function as competitors that result in a mutual restriction in their chromatin site occupancy when they are present together.

### 3.4.5 Sequence Analysis of ER $\alpha$ vs. ER $\beta$ Binding Sites

We next examined whether the genomic sequences to which both ER $\alpha$  and ER $\beta$  bind contain a recognizable estrogen response element (ERE) motif by performing a DNA-binding motif search. We considered a 13-base-pair site with up to two positions varying from the canonical ERE (GGTCAnnnTGACC) as a putative ERE motif (full-ERE). Among the ER $\alpha$ [ $\alpha$  cells] binding regions (Figure 3.4A), 63% contained full-ERE sequences, 23% had ERE half-sites, and 14% had no ERE-like sequences. The ERE motif distribution in the ER $\beta$ [ $\beta$  cells] binding regions (Figure 3.4B) was very similar to that of the ER $\alpha$ [ $\alpha$  cells] regions. These results support the idea that ER $\alpha$  and ER $\beta$ , in the absence of their respective ER subtype partners, tend to bind predominantly to full EREs.

To further address whether ER $\alpha$  and ER $\beta$  will bind to ERE-like sequences when they are present together in the cells, we performed the ERE motif search on the ER binding sites in cells expressing both ER $\alpha$  and ER $\beta$ . ER $\alpha$ [ $\alpha\beta$  cells] binding sites have an ERE motif distribution (Figure 3.4C) resembling the distribution of the ER $\alpha$ [ $\alpha$  cells] sites and ER $\beta$ [ $\beta$  cells] sites (Figure 3.4A and 3.4B, respectively). Surprisingly, we found that the ER $\beta$ [ $\alpha\beta$  cells] binding regions (Figure 3.4D) contained a lower percentage of ERE sequences as compared to ER $\beta$ [ $\beta$  cells] binding regions (Figure 3.4B). In fact, almost one-third of the ER $\beta$ [ $\alpha\beta$  cells] binding site regions do not contain any ERE-like (full and half ERE) sequences.

These observations are a further indication that when both receptors are present in the same cells, there is mutual interaction between ER $\alpha$  and ER $\beta$  that affects their chromatin binding site preferences, with ER $\alpha$  being dominant. Interestingly, ER $\alpha$  dominance over ER $\beta$  in chromatin binding is most pronounced at ERE-containing sites. This suggests that there is a preference for formation of ER $\alpha$  homodimers that drive ER $\beta$  in  $\alpha\beta$  cells to occupy sites less enriched in EREs.

There is also an alternative possibility, namely that instead of ER $\beta$  homodimers being forced to occupy new binding sites by the presence of ER $\alpha$  homodimers, ER $\alpha$ /ER $\beta$  heterodimers might bind preferentially to alternate binding sites, ones that are not occupied by either ER homodimer. This possibility could have interesting consequences on the transcriptional output when ER $\alpha$  and ER $\beta$  are both present in cells.

### 3.4.6 Correlation Between ER Binding Sites in $\alpha\beta$ Cells and E<sub>2</sub>-Regulated Genes

Having shown the competitive nature of ER $\alpha$  and ER $\beta$  recruitment to the chromatin binding sites in cells expressing both receptors, we next investigated the association between ER $\alpha$  and ER $\beta$  recruitment to cis-regulatory sites and E<sub>2</sub>-mediated transcriptional responses in ER $\alpha\beta$  cells. This was accomplished by comparing the promoter regions (10kb upstream and 10kb downstream of the transcription start site (TSS)) of 90 genes that are either E<sub>2</sub>-stimulated or repressed in these cells [3] and that have at least one site bound specifically by either receptor alone (ER $\alpha$  unique sites or ER $\beta$  unique sites) or sites shared by both receptors (ER $\alpha$ /ER $\beta$  shared sites) (Figure 3.5A).

Our analysis showed that the promoter regions of E<sub>2</sub>-repressed genes in  $\alpha\beta$  cells were three times more likely to have binding sites unique to ER $\beta$  than to ER $\alpha$ . This suggests that ER $\beta$  recruited more strongly than ER $\alpha$  to the promoter regions of E<sub>2</sub>-repressed genes. Intriguingly, for E<sub>2</sub>-stimulated genes, promoter regions were highly (2 fold) associated with binding sites that are shared by both ER $\alpha$  and ER $\beta$ . These binding sites can be occupied either by an ER $\alpha$ /ER $\beta$  heterodimer complex or occupied at different times by ER $\alpha$  or ER $\beta$  homodimers.

A number of studies have previously examined the transcriptional activities of ER $\beta$  and ER $\alpha$  in breast cancer cells [3, 8], but as far as we know, the results presented here are the first to associate ER upstream genome target binding to the transcriptional output of the corresponding regulated genes. Although we have shown earlier, on a genome-wide scale, that ER $\beta$  may be weaker in competing with ER $\alpha$  for binding sites, this study suggests that in the promoter regions of some E<sub>2</sub>-repressed genes, ER $\beta$  can successfully compete with ER $\alpha$  and exclude it from binding.

To further characterize ER $\alpha$  and ER $\beta$  functional mechanisms and possible co-modulatory effects on gene regulation, we monitored ER $\alpha$  and ER $\beta$  recruitment to chromatin target sites both upstream and downstream of the transcription start site of the well known E<sub>2</sub>-regulated gene, FOS, in the three cell-types after E<sub>2</sub> treatment. We first measured the transcript level of FOS by quantitative RT-PCR in response to E<sub>2</sub> treatment in the three cell types (Figure 3.5B). In both ER $\alpha$ -only and  $\alpha\beta$  cells, the E<sub>2</sub>-stimulated expression of FOS was very similar; however, we saw a reduced (ca. 40%) expression of FOS in cells expressing only ER $\beta$ . This suggested that ER $\beta$  might be a weaker transcriptional activator of this gene relative to ER $\alpha$ .

We then examined both ER $\alpha$  and ER $\beta$  recruitment to three potential ER binding sites (identified by our ChIP-chip data) by ChIP-qPCR. The sites are denoted as FOS\_enh1, FOS\_enh2 and FOS\_3'end. The first two sites are located ~20kb upstream, whereas the third is located ~5kb downstream of the FOS TSS. The binding data is shown in Figure 3.5C. We observed that ER $\alpha$  and ER $\beta$  could bind to the FOS\_enh2 site in all three types of cells and was not affected by the presence of its ER subtype partner; however, FOS\_enh1 and FOS\_3'end sites were bound exclusively by ER $\alpha$ . We hypothesize that either one or both of these sites



(FOS\_enh1 and FOS\_3'end) are responsible for the enhanced transcription of FOS seen in ER $\alpha$ -only and in ER  $\alpha\beta$  cells.

This hypothesis was tested by comparing PolII ChIP-chip data published by Carroll et al. [21] with the ER binding sites around FOS. PolII binding was detected at three separate sites in the vicinity of the FOS TSS. Interestingly, we found that one of the PolII sites was located at FOS\_enh1 (ER $\alpha$ -exclusive binding site). The PolII data suggests that the enhanced FOS transcript level in ER $\alpha$ -only and ER $\alpha\beta$  cells might be due to the recruitment of additional PolII by ER $\alpha$  to the FOS\_enh1 site, which might then bring the recruited PolII to the TSS of FOS, perhaps by a looping mechanism [33]. This would be consistent with the reduced FOS expression in ER $\beta$ -only cells, because ER $\beta$  does not access the FOS\_enh1 site, so no additional PolII would be brought to the TSS of FOS by ER $\beta$ .

#### 3.4.7 Binding Site Distribution with the ER $\alpha$ -Selective Ligand (PPT) vs. E<sub>2</sub>

To assess the effect of the ER $\alpha$ -selective ligand, PPT, on ER $\alpha$  chromatin binding, we examined binding site occupancy after PPT exposure in ER $\alpha$ -only or ER $\alpha\beta$  cells (Table 3.1). We found that the ER $\alpha$  binding sites, after PPT treatment, were very similar to those observed after E<sub>2</sub> treatment (Figure 3.6A and 3.6B). In ER $\alpha$ -only cells, 92% of ER $\alpha$  binding sites after PPT treatment were also found in the E<sub>2</sub> treatment group, and in ER $\alpha\beta$  cells, 83% of ER $\alpha$  binding sites after PPT treatment were the same as the ER $\alpha$  binding sites from the E<sub>2</sub> treatment group. The data suggests that the ER $\alpha$ -PPT complex formed is quite similar to the ER $\alpha$ -E<sub>2</sub> complex in terms of enabling ER $\alpha$  to access the chromatin binding sites to which ER $\alpha$  binds in both ER $\alpha$ -cells and ER $\alpha\beta$ -cells.

In contrast to E<sub>2</sub>, PPT has extremely low affinity for ER $\beta$ , less than 0.2% that of E<sub>2</sub> [13]; therefore, it is very unlikely to bind to or activate ER $\beta$  in cells containing both ERs. Therefore, we hypothesized that in ER $\alpha\beta$ -cells, the unliganded ER $\beta$  would be unable to compete effectively against the ER $\alpha$ -PPT complex for ER binding sites. Consistent with this hypothesis was the ChIP-chip data showing that the number of ER $\alpha$ [ $\alpha\beta$  cells] binding sites increases from 3252 to 3466 binding sites when these cells are treated with PPT as compared to E<sub>2</sub>. This slight increase in ER $\alpha$ [ $\alpha\beta$  cells] binding sites after PPT treatment is consistent with the concept that activated ER $\alpha$  and ER $\beta$  act as competitors when they are present in the same cells, but that unoccupied ER $\beta$  is less effective as a competitor of ER $\alpha$  for chromatin binding.

#### 3.4.8 Binding Site Distribution with the ER $\beta$ -Selective Ligand (ERB-041) vs. E<sub>2</sub>

In addition to our ChIP-chip analysis of ER $\alpha$  chromatin binding after PPT treatment, we also characterized the chromatin binding of ER $\beta$  after treatment with ERB-041, an ER $\beta$ -selective ligand [11], in MCF-7 cells containing either ER $\beta$ -only or both ER $\alpha$  and ER $\beta$ . We found that when treated with ERB-041, ER $\beta$  bound to 1042 sites in ER $\beta$ -only cells and 1109 sites in ER $\alpha\beta$  cells (Table 3.1). After ERB-041 treatment, ER $\beta$  bound to many ER sites different from those following E<sub>2</sub> treatment. In ER $\beta$ -cells after ERB-041 treatment, only 51% of the ER $\beta$  binding sites overlapped the ER $\beta$  binding sites in the E<sub>2</sub> treatment group (Figure 3.6C). Similarly, in ER $\alpha\beta$ -cells, only 47% of the binding sites bound by ER $\beta$  after ERB-041 treatment overlapped the sites bound by ER $\beta$  after E<sub>2</sub> treatment (Figure 3.6D).

These results provide an indication of the complexity in relating the conformation and activation of the different ER-ligand complexes to their recruitment at chromatin target sites. The ER $\alpha$ -PPT complex behaves in a manner much like that of the ER $\alpha$ -E<sub>2</sub> complex, binding to many

of the same sites, and the addition of ER $\beta$  (occupied in the presence of E<sub>2</sub>, but unoccupied in the presence of PPT) has relatively little effect on ER $\alpha$  binding (Figure 3.6A and 3.6B). In contrast, the ER $\beta$ -ERB-041 binding sites are often not the same as those of the ER $\beta$ -E<sub>2</sub> complex. While this suggests that the conformations induced in ER $\beta$  by the binding of ERB-041 vs. E<sub>2</sub> might be different, this is not apparent from a comparison of their X-ray crystal (PDB entry 2j7x for ER $\beta$  with E<sub>2</sub> (unpublished) and 1x7b for ER $\beta$  with ERB-041 [28]).

### 3.4.9 Overlap of ER $\alpha$ and ER $\beta$ Binding Sites with Subtype-Selective Ligands

After characterizing the patterns of chromatin binding by ER $\alpha$  and ER $\beta$  when liganded by their corresponding subtype-selective ligands vs. E<sub>2</sub>, we further examined the relationship between ER $\alpha$ -PPT and ER $\beta$ -ERB-041 chromatin binding sites. First, we compared ER $\alpha$  [ $\alpha$  cells] binding sites with ER $\beta$  [ $\beta$  cells] binding sites (Figure 3.7A). Previously, we had defined the overlap of binding sites in this case as sites “in common” to ER $\alpha$  or ER $\beta$ , meaning that these sites could be occupied by either ER subtype, provided that the other subtype was not present. Curiously, there were fewer ER $\alpha$ -PPT [ $\alpha$ -cells] and ER $\beta$ -ERB-041 [ $\beta$ -cells] sites in common (509) than were in common when ER $\alpha$  and ER $\beta$  were bound with E<sub>2</sub> (1386, Figure 3.3A). This suggests that the conformation induced by the binding of a subtype-selective ligand to its respective ER subtype is not precisely the same as that induced by the binding of E<sub>2</sub>, and as a result, ER $\alpha$  and ER $\beta$  bound by their respective subtype-specific ligands have more distinct binding preferences than do the E<sub>2</sub>-bound ER subtypes, meaning fewer sites in common.

We next investigated the relationship between ER $\alpha$  [ $\alpha\beta$  cells] binding sites with ER $\beta$  [ $\alpha\beta$  cells] binding sites (Figure 3.7B). These were sites previously defined as “shared” by ER $\alpha$  and ER $\beta$ , meaning that either one or both of the receptors could bind to these sites when both were present

in cells (Figure 3.3B). Interestingly, in ER $\alpha\beta$  cells, when one or the other of the ER subtype was occupied by its respective subtype-specific ligand, there were, in fact, a larger number of sites shared by ER $\alpha$  and ER $\beta$  (730) than when both ERs were co-occupied by E<sub>2</sub> (579) (compare Figure 3.3B and Figure 3.7B). Because PPT and ERB-041 are ER subtype-selective ligands, when ER $\alpha\beta$  cells are treated with one or the other of these ligands, one of the ER subtypes will be liganded but the other will be unliganded. The greater number of shared sites when the subtype-selective ligands are used, indicates that unoccupied ERs are less able to compete with occupied ERs for the same binding sites. This is consistent with our previous data showing that both ERs compete with one another quite well when they are present together in the cells following E<sub>2</sub> treatment, in which case, both ERs will be occupied by E<sub>2</sub> because it binds well to both ER $\alpha$  and ER $\beta$ .

### 3.5 DISCUSSION

Previous studies done by our group and others [3, 4, 25, 34] have shown that ER $\beta$  has a significant impact in modulating the expression of genes regulated by ER $\alpha$  in breast cancer cells. However, understanding this orchestrated transcriptional program by both ER $\alpha$  and ER $\beta$  will not be complete without the examination and identification of the chromatin targets of ER $\alpha$  and ER $\beta$  when they are either separately or present together in breast cancer cells under various treatment conditions. Therefore, in this study, we seek to identify, in a genome-wide level, the chromatin localizations of ER $\alpha$  and ER $\beta$  in MCF-7 breast cancer cells and examine how ER selection of chromatin binding sites is affected by their state of ligand occupancy.

When both receptors are coexpressed in the cells, we observed that the number of binding sites that can be shared by both ER $\alpha$  and ER $\beta$  is very much reduced (579 sites, Figure 3.3B)

compared to the number of sites that are common (1386 sites, Figure 3.3A). This result suggests that, on a genome scale, there is active competition between the two receptors for chromatin binding sites when they are present together in cells and they can no longer bind unobstructed to their native binding sites as in the case when they are present alone in the cells, therefore ending up sharing a much limited set of binding sites. In order to examine the how the presence of their subtype partner has on ER binding activities, we assessed the binding profile of ER $\alpha$  and ER $\beta$  in the copresence or absence of its isoform partner. We found that 92% of ER $\alpha$ [ $\alpha\beta$  cells] binding sites overlap ER $\alpha$ [ $\alpha$  cells] binding sites and 40% of ER $\beta$ [ $\alpha\beta$  cells] binding sites overlap ER $\beta$ [ $\beta$  cells] binding sites. Like in all competitive interactions, there will be a dominant player and our results indicate that ER $\alpha$  is the more dominant isoform as it appears that the presence of ER $\beta$  has a small influence on ER $\alpha$  binding sites selections since ER $\alpha$ [ $\alpha\beta$  cells] binds to most of its native sites (ER $\alpha$ [ $\alpha$  cells]) despite the presence of liganded ER $\beta$ . But the presence of liganded ER $\alpha$  has a greater influence on perturbing ER $\beta$  chromatin binding site profile with ER $\beta$ [ $\alpha\beta$  cells] binding to new alternative binding sites differ from their native binding sites (ER $\beta$ [ $\beta$  cells]). Overall, from a systemic point of view, our results exemplify the competitive nature of ER $\alpha$  and ER $\beta$  in chromatin binding with ER $\alpha$  being more “dominant” in competing for ER binding sites and also the binding dynamism of the two ER subtypes with ER $\beta$  binding to new alternative sites less enriched in EREs when it native binding sites were bound by ER $\alpha$ .

A number of studies have previously examined the transcriptomic activities of ER $\alpha$  and ER $\beta$  with the bulk of evidence implying that ER $\beta$  has growth-suppressive activities [5-7]. Our ChIP-chip analyses in ER $\alpha\beta$  cells have revealed new mechanistic insight into ER $\beta$  modulated gene expression. Our analysis showed that the promoter regions of certain E2 repressed genes were

three times more likely to have binding sites bound exclusively by ER $\beta$ . Although ER $\alpha$  may be the dominant ER isoform, but intriguingly, within the promoter regions of these E2-repressed genes, ER $\beta$  is able to compete and exclude ER $\alpha$  from binding. This transcriptional signaling activity may be, at least in part, responsible for ER $\beta$  growth-suppressive activities. The results of our sequence analysis (Figure 3.4) revealed that both ER $\alpha$ [ $\alpha$  cells] and ER $\beta$ [ $\beta$  cells] are binding mostly to chromatin targets containing ERE motifs. This observation fits well with the current consensus that ER $\alpha$  and ER $\beta$  both appear to recognize the same estrogen response element motif. The interesting observation is that although the ERE motif distribution of ER $\alpha$ [ $\alpha\beta$  cells] binding regions is very similar to distributions found in both ER $\alpha$ [ $\alpha$  cells] and ER $\beta$ [ $\beta$  cells] binding regions but ER $\beta$ [ $\alpha\beta$  cells] is binding to sites which contained a lower percentage of ERE sequences. This is likely the reflection of the competitive nature of the two ER isoforms when they are coexpressed in the same cells; ER $\beta$ [ $\alpha\beta$  cells] being denied of its native sites by the dominant liganded ER $\alpha$  will have to occupy alternative binding sites less enriched with EREs.

We also examined the impact of subtype-selective ligands on the chromatin binding profiles of ER $\alpha$  and ER $\beta$ . Our Chip-chip analysis showed that most of the sites bound by ER $\alpha$ [ $\alpha$  cells] and ER $\alpha$ [ $\alpha\beta$  cells] after PPT treatment were also found in the E<sub>2</sub> treatment group (Figure 3.6A and 3.6B), suggesting that the conformation change of the ER $\alpha$ -PPT complex should be quite similar to the ER $\alpha$ -E<sub>2</sub> complex thus enabling similar access to the chromatin. However, we noted that the sites bound by ER $\beta$ [ $\beta$  cells] and ER $\beta$ [ $\alpha\beta$  cells] after ERB-041 treatment are quite distinct from the binding sites in the E<sub>2</sub> treatment group (Figure 3.6C and 3.6D). Only approximately half of the ER $\beta$  binding sites are shared between the two treatment groups indicating that the conformations induced on the ER $\beta$  complexes by E<sub>2</sub> and ERB-041 might be

different. Intriguing, a comparison of their X-ray crystal shows that the ER $\beta$  conformations by the two ligands are very similar suggesting that it might be due to the differences in recruiting certain cofactors by the different liganded ER $\beta$  complexes, which may determine chromatin accessibility. Altogether, our data reveal the additional complexity in relating the conformation of the different ER-ligand complexes to their recruitment to chromatin in vivo. Our analysis showed that the number of shared sites (either one or both of the receptors can bind in ER $\alpha\beta$  cells) increased when the ER subtype was occupied by its respective subtype-specific ligand compared to both ERs co-occupied by E2 (Figure 3.7B v.s. Figure 3.3B). The increase in number of shared sites when the subtype-selective ligands are used indicates that it is the liganded ER subtype that principally determines the location of the chromatin binding sites.

There is abundant evidence that in addition to ER $\alpha$ , ER $\beta$  plays a critical role in human breast cancer as ER-positive human breast cancers usually contain both ER $\alpha$  and ER $\beta$ . Our ChIP-chip analysis of both ERs provides us with genome scale snapshots of ERs' cis-regulatory elements which can help us elucidate the activities of both ERs and how each ER subtype will impact the binding profile of the other subtype in breast cancer. Our work suggests that when both ER subtypes are present in the cells, they each restrict the other ER from accessing their preferred binding sites but overall, ER $\alpha$  appears to be the dominating ER subtype. Our studies also indicate that there is a tendency for both ER $\alpha$  and ER $\beta$  to bind to sites enriched with EREs. Our data further define on a genome scale the chromatin binding sites of ER $\alpha$  and ER $\beta$  under subtype-selective ligands treatment. The results point to an important role of the liganded ER; that it is principally responsible for determining the selection of chromatin binding sites. Taken together, our ChIP-chip studies provide insights to the dynamic interplay between ER $\alpha$  and ER $\beta$  in selecting and competing for chromatin binding sites in breast cancer cells.

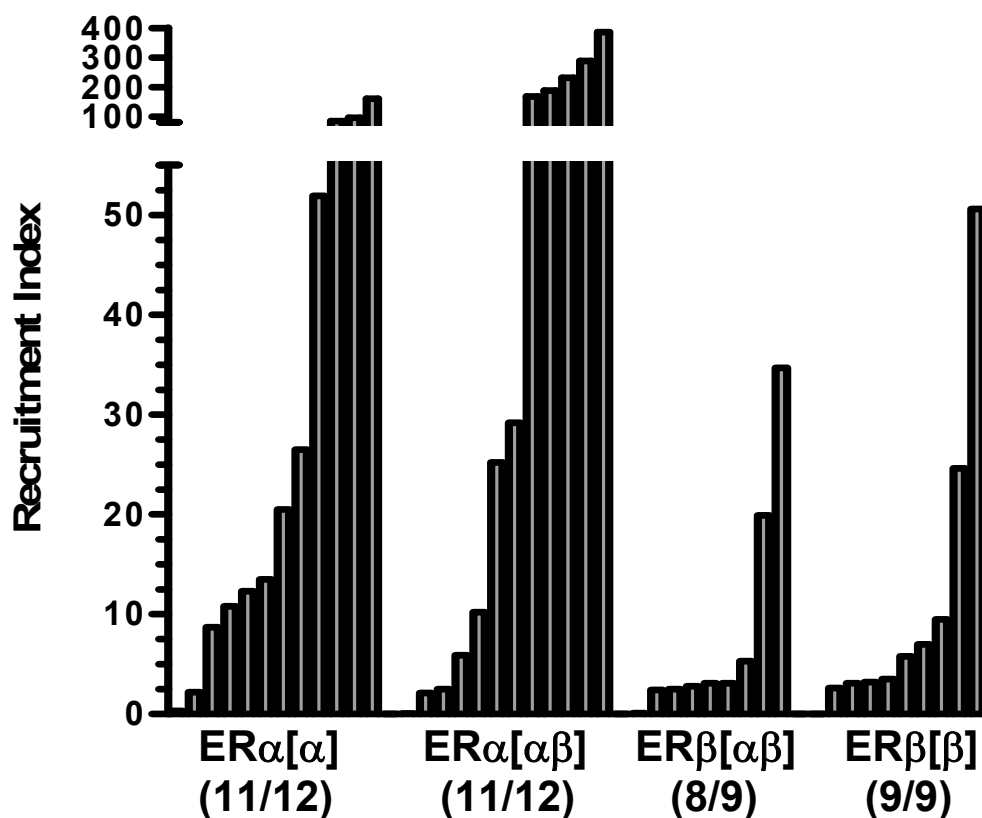
### 3.6 TABLE AND FIGURES

**Table 3.1.** Summary of ER binding sites in the three MCF-7 cells expressing ER $\alpha$  only, both ER $\alpha$  and ER $\beta$ , or ER $\beta$  only.

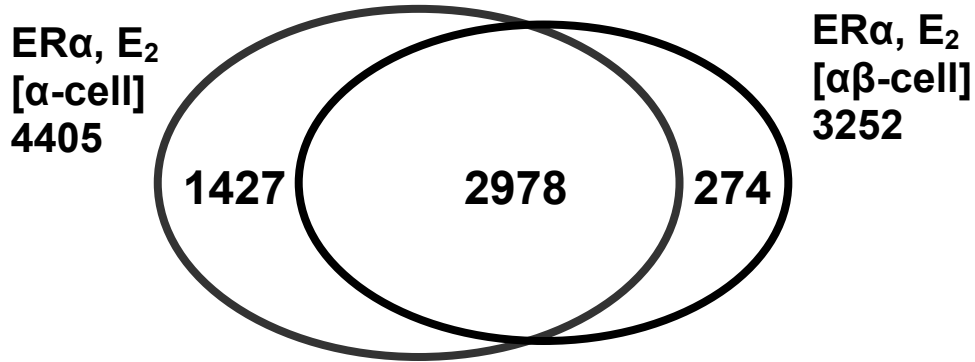
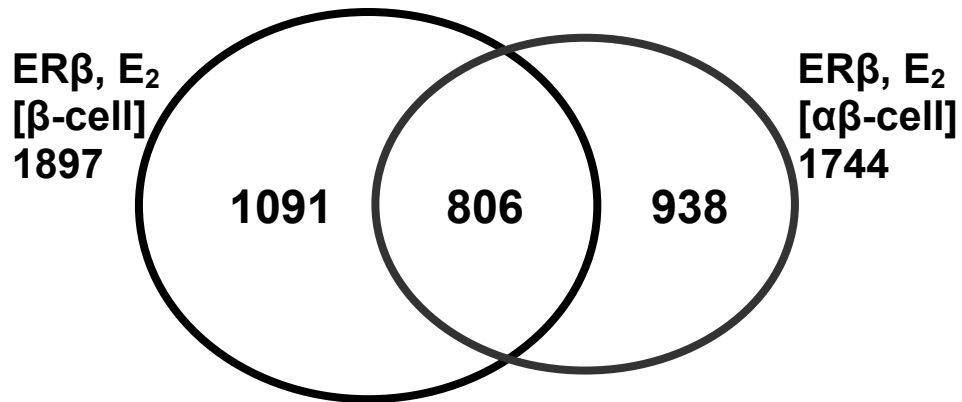
	<b>Cell Type</b>	<b>Ligand</b>	<b>Antibodies</b>	<b>Binding sites*</b>
1	MCF7 [ $\alpha$ cells]	E2	anti-ER $\alpha$	4405
2	MCF7 [ $\alpha$ cells]	PPT	anti-ER $\alpha$	3269
3	MCF7 [ $\alpha\beta$ cells]	E2	anti-ER $\alpha$	3252
4	MCF7 [ $\alpha\beta$ cells]	E2	anti-ER $\beta$	1744
5	MCF7 [ $\alpha\beta$ cells]	PPT	anti-ER $\alpha$	3466
6	MCF7 [ $\alpha\beta$ cells]	ERB-041	anti-ER $\beta$	1109
7	MCF7 [ $\beta$ cells]	E2	anti-ER $\beta$	1897
8	MCF7 [ $\beta$ cells]	ERB-041	anti-ER $\beta$	1042

\*Values are the mean from three independent experiments, with each experiment done in duplicate.

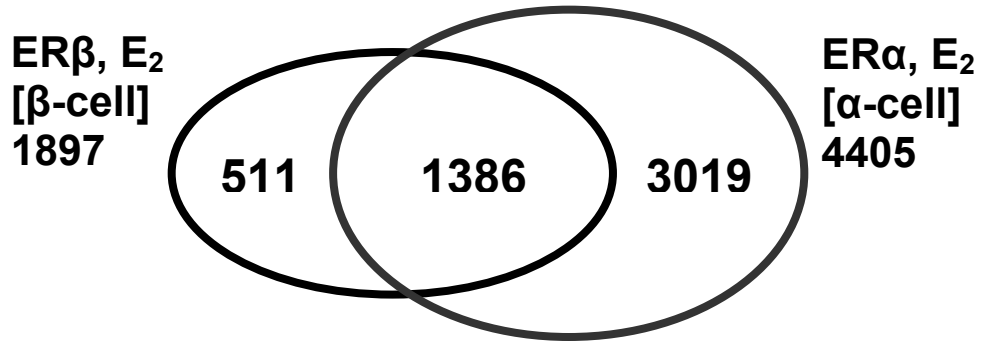
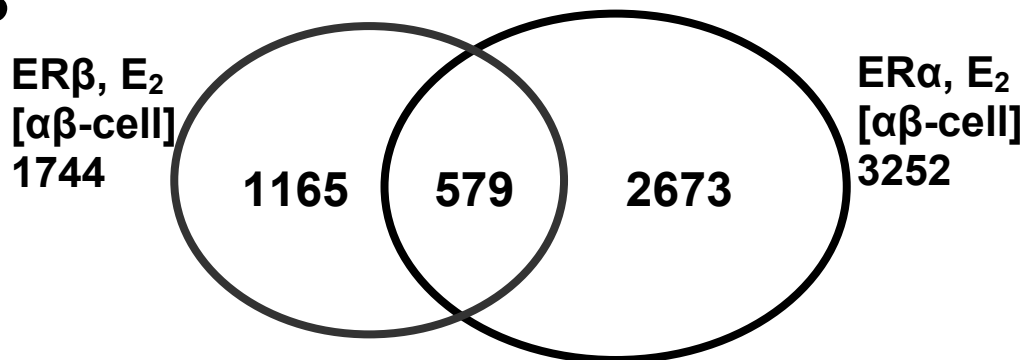




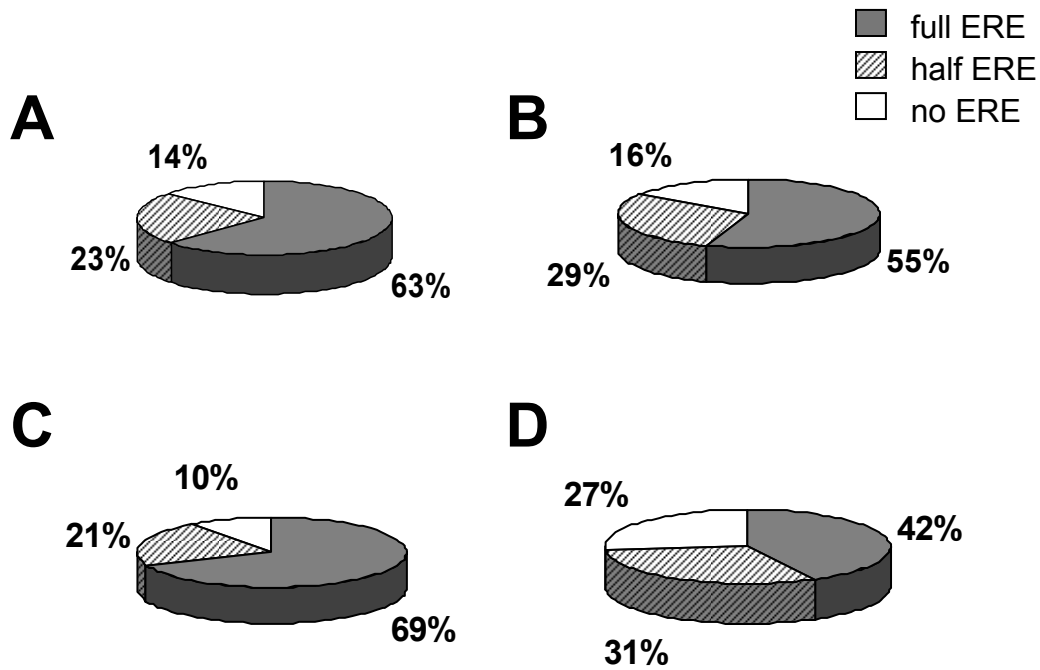
**Figure 3.1.** Validation of ChIP-chip binding sites. ER $\alpha$  and ER $\beta$  chromatin binding (by conventional ChIP assays) were measured by quantitative PCR after 45min of E<sub>2</sub> treatment of MCF-7 cells differentially expressing ER $\alpha$  and/or ER $\beta$ . Data are expressed as recruitment index (ER ChIP/Input) and is the average of two replicates. Binding sites are considered validated if the recruitment index is >2. ChIP-qPCR results are shown for (from left to right) ER $\alpha$  recruitment in ER $\alpha$ -only cells, ER $\alpha$  recruitment in cells containing both ERs, ER $\beta$  recruitment in cells containing both ERs, and ER $\beta$  recruitment in ER $\beta$ -only cells.

**A****B**

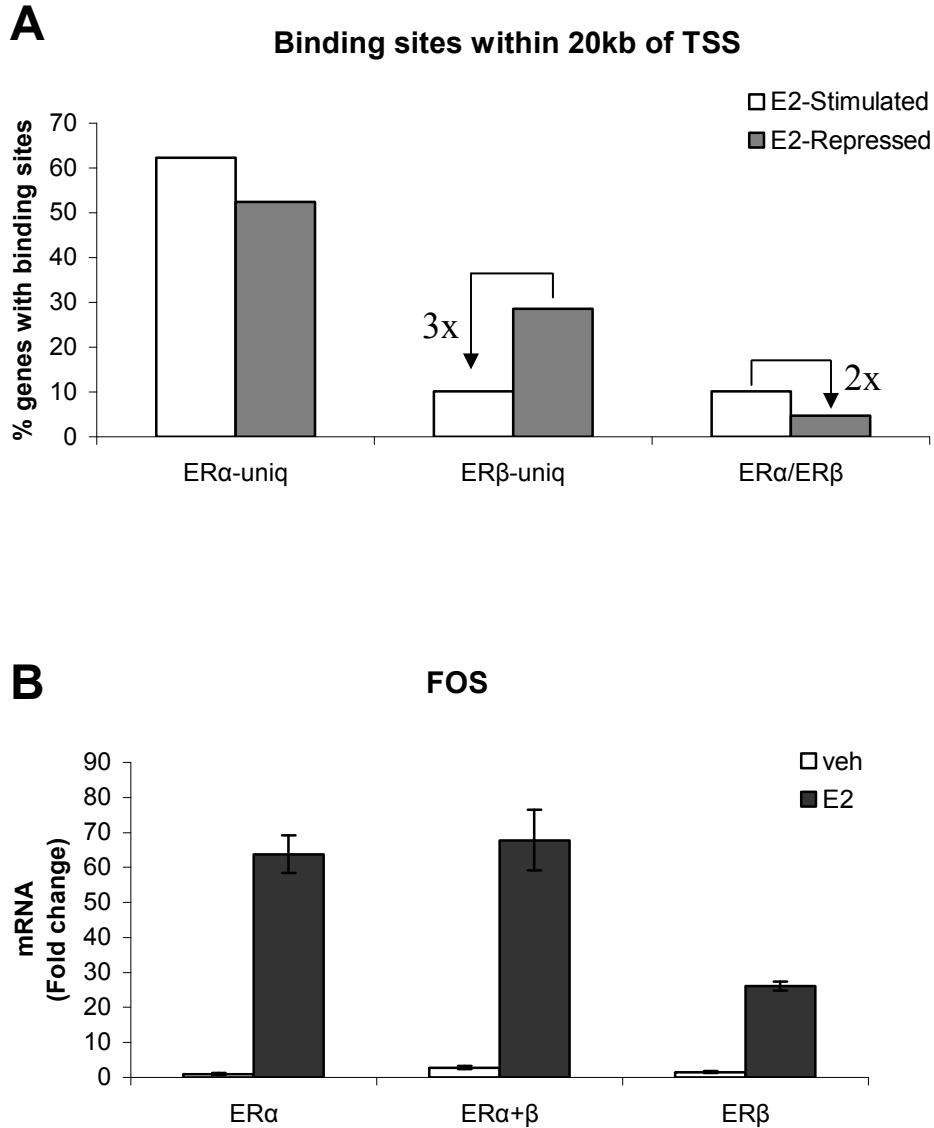
**Figure 3.2.** Effect of ER subtype partner on ER binding site distribution with E<sub>2</sub> treatment. (A) The introduction of ERβ into the cells has a relatively minor effect on the distribution of ERα binding sites. (B) ERα has a more pronounced effect on the distribution of ERβ binding sites.

**A****B**

**Figure 3.3.** Venn diagrams comparing the occupancy of ER binding sites by ERα and ERβ when they are present either separately or together in cells treated with E<sub>2</sub> (A) ERα or ERβ can each occupy many of the same sites when the other ER subtype is not present in the cells. (B) When both receptors are present in the cells, ERα and ERβ share a more limited number of sites.

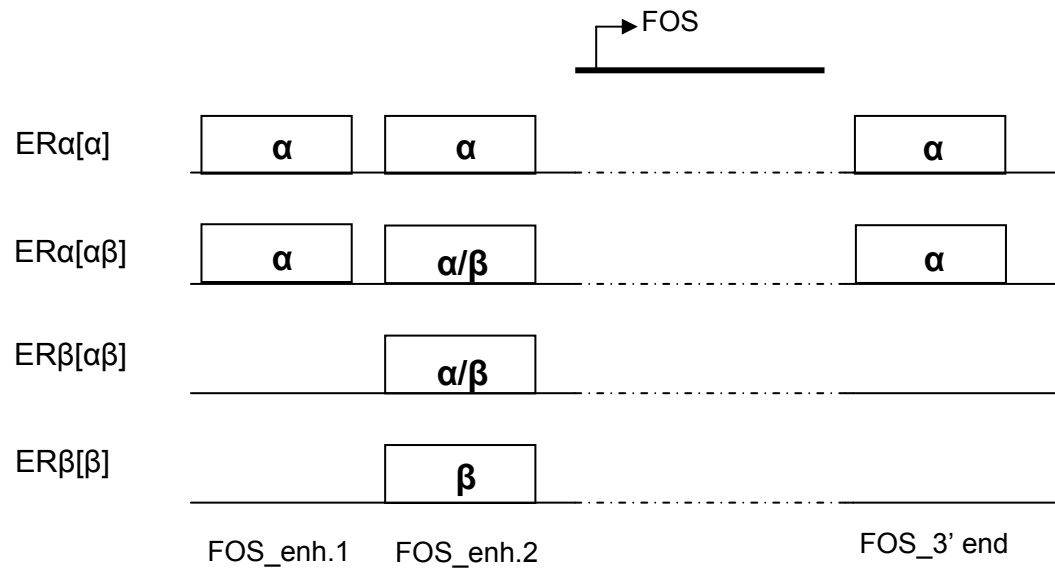


**Figure 3.4.** Presence of ERE sequences in ERα or ERβ binding sites with E<sub>2</sub> treatment. Binding sites were probed for the presence of full ERE (allowing for up to two base deviations from the canonical ERE), half ERE, and no ERE motifs. (A) ERα binding sites in ERα-only cells. (B) ERβ binding sites in ERβ-only cells. (C) ERα binding sites in cells containing both ERα and ERβ receptors. (D) ERβ binding sites in cells containing both ERα and ERβ receptors.

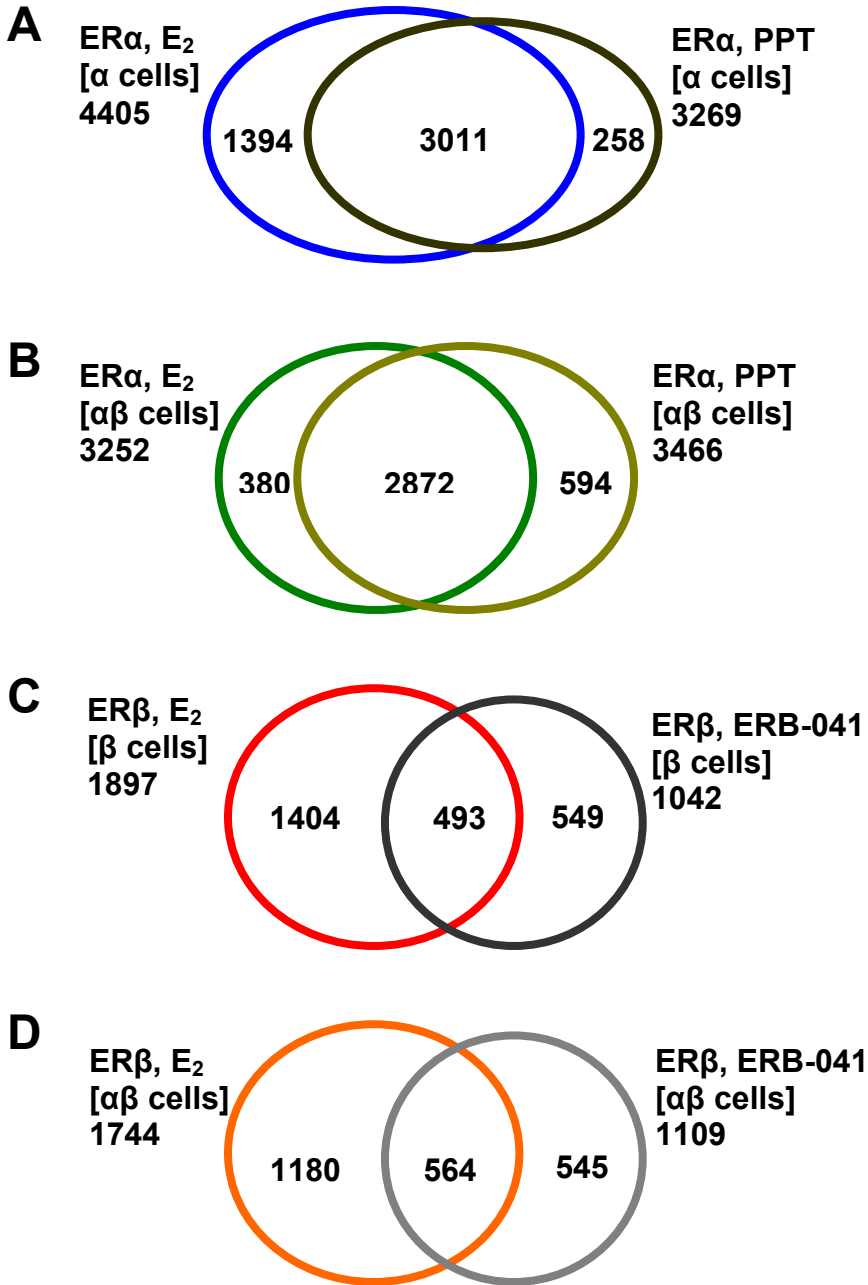


**Figure 3.5.** Correlation between ER binding and transcriptional output in response to E<sub>2</sub>. (A) Correlation between E<sub>2</sub>-regulated genes and binding of ERα-unique (only ERα binds), ERβ-unique (only ERβ binds), or ERα/ERβ (the sites are shared by both ERs) within  $\pm 10$ kb of the TSS of the genes. (B) FOS mRNA levels were assessed by quantitative PCR after 4hr treatment of MCF-7 cells differentially expressing ERα and/or ERβ. Data represent average fold change  $\pm$  SD for three independent replicates. (C) ERα and ERβ chromatin binding (by conventional ChIP assays) were measured by quantitative PCR after 45min E<sub>2</sub> treatment of MCF-7 cells expressing ERα and/or ERβ. ERα and ERβ occupancy of three different ER binding sites that are closest to the FOS gene are presented graphically. Enh., denotes enhancer.

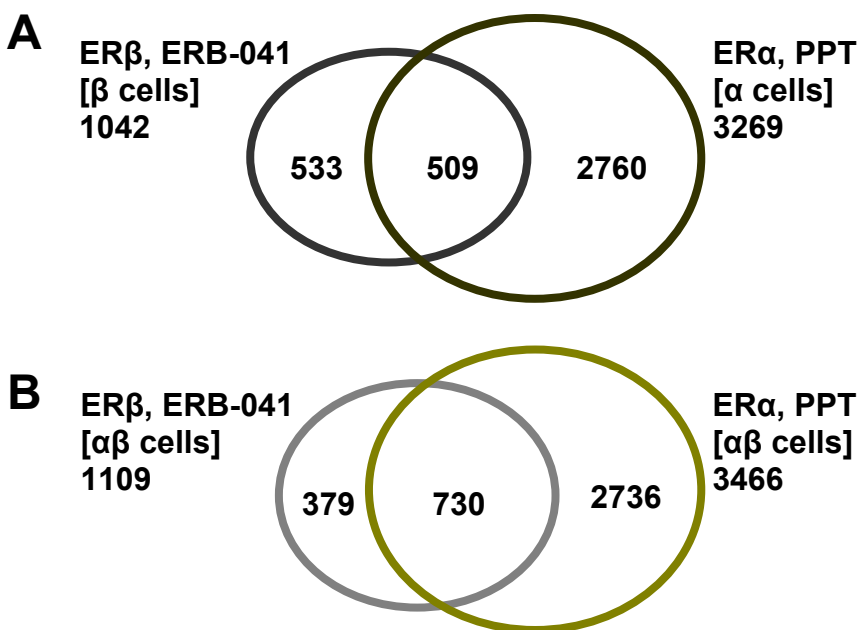
**C**



**Figure 3.5.** (cont.)



**Figure 3.6.** Venn diagrams comparing ER binding site occupancy after cell treatment with the ER $\alpha$ -selective ligand (PPT) or ER $\beta$ -selective ligand (ERB-041) vs. E<sub>2</sub>. (A) ER $\alpha$  binding sites in ER $\alpha$ -only cells (E<sub>2</sub> vs. PPT treatment). (B) ER $\alpha$  binding sites in cells containing both ER $\alpha$  and ER $\beta$  (E<sub>2</sub> vs. PPT treatment). (C) ER $\beta$  binding sites in ER $\beta$ -only cells (E<sub>2</sub> vs. ERB-041 treatment). (D) ER $\beta$  binding sites in cells containing both ER $\alpha$  and ER $\beta$  (E<sub>2</sub> vs. ERB-041 treatment).



**Figure 3.7.** Venn diagrams showing the overlap of ER $\alpha$  binding sites (PPT treatment) and ER $\beta$  binding sites (ERB-041 treatment). (A) ER $\beta$  binding sites in ER $\beta$ -only cells (ERB-041 treatment) and ER $\alpha$  binding sites in ER $\alpha$ -only cells (PPT treatment). (B) ER $\beta$  and ER $\alpha$  binding sites in cells containing both ER $\alpha$  and ER $\beta$  (ERB-041 vs. PPT treatment).



### 3.7 REFERENCES

1. Deroo, B.J. and K.S. Korach, *Estrogen receptors and human disease*. J Clin Invest, 2006. **116**(3): p. 561-70.
2. Nilsson, S., et al., *Mechanisms of estrogen action*. Physiol Rev, 2001. **81**(4): p. 1535-65.
3. Chang, E.C., et al., *Impact of estrogen receptor beta on gene networks regulated by estrogen receptor alpha in breast cancer cells*. Endocrinology, 2006. **147**(10): p. 4831-42.
4. Chang, E.C., et al., *Estrogen Receptors alpha and beta as determinants of gene expression: influence of ligand, dose, and chromatin binding*. Mol Endocrinol, 2008. **22**(5): p. 1032-43.
5. Frasor, J., et al., *Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype*. Endocrinology, 2003. **144**(10): p. 4562-74.
6. Paruthiyil, S., et al., *Estrogen receptor beta inhibits human breast cancer cell proliferation and tumor formation by causing a G2 cell cycle arrest*. Cancer Res, 2004. **64**(1): p. 423-8.
7. Strom, A., et al., *Estrogen receptor beta inhibits 17beta-estradiol-stimulated proliferation of the breast cancer cell line T47D*. Proc Natl Acad Sci U S A, 2004. **101**(6): p. 1566-71.
8. Williams, C., et al., *A genome-wide study of the repressive effects of estrogen receptor beta on estrogen receptor alpha signaling in breast cancer cells*. Oncogene, 2008. **27**(7): p. 1019-32.
9. Frasor, J., et al., *Gene expression preferentially regulated by tamoxifen in breast cancer cells and correlations with clinical outcome*. Cancer Res, 2006. **66**(14): p. 7334-40.
10. Kraichely, D.M., et al., *Conformational changes and coactivator recruitment by novel ligands for estrogen receptor-alpha and estrogen receptor-beta: correlations with biological character and distinct differences among SRC coactivator family members*. Endocrinology, 2000. **141**(10): p. 3534-45.
11. Malamas, M.S., et al., *Design and synthesis of aryl diphenolic azoles as potent and selective estrogen receptor-beta ligands*. J Med Chem, 2004. **47**(21): p. 5021-40.
12. Meyers, M.J., et al., *Estrogen receptor subtype-selective ligands: asymmetric synthesis and biological evaluation of cis- and trans-5,11-dialkyl- 5,6,11, 12-tetrahydrochrysenes*. J Med Chem, 1999. **42**(13): p. 2456-68.
13. Stauffer, S.R., et al., *Pyrazole ligands: structure-affinity/activity relationships and estrogen receptor-alpha-selective agonists*. J Med Chem, 2000. **43**(26): p. 4934-47.
14. Katzenellenbogen, B.S. and J. Frasor, *Therapeutic targeting in the estrogen receptor hormonal pathway*. Semin Oncol, 2004. **31**(1 Suppl 3): p. 28-38.
15. Kurebayashi, J., et al., *Expression levels of estrogen receptor-alpha, estrogen receptor-beta, coactivators, and corepressors in breast cancer*. Clin Cancer Res, 2000. **6**(2): p. 512-8.
16. Saji, S., M. Hirose, and M. Toi, *Clinical significance of estrogen receptor beta in breast cancer*. Cancer Chemother Pharmacol, 2005. **56 Suppl 1**: p. 21-6.
17. Shaaban, A.M., et al., *Declining estrogen receptor-beta expression defines malignant progression of human breast neoplasia*. Am J Surg Pathol, 2003. **27**(12): p. 1502-12.

18. Speirs, V., et al., *Oestrogen receptor beta: what it means for patients with breast cancer*. Lancet Oncol, 2004. **5**(3): p. 174-81.
19. Palmieri, C., et al., *Estrogen receptor beta in breast cancer*. Endocr Relat Cancer, 2002. **9**(1): p. 1-13.
20. Roger, P., et al., *Decreased expression of estrogen receptor beta protein in proliferative preinvasive mammary tumors*. Cancer Res, 2001. **61**(6): p. 2537-41.
21. Carroll, J.S., et al., *Genome-wide analysis of estrogen receptor binding sites*. Nat Genet, 2006. **38**(11): p. 1289-97.
22. Lin, C.Y., et al., *Whole-genome cartography of estrogen receptor alpha binding sites*. PLoS Genet, 2007. **3**(6): p. e87.
23. Kininis, M., et al., *Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters*. Mol Cell Biol, 2007. **27**(14): p. 5090-104.
24. Liu, Y., et al., *The genome landscape of ERalpha- and ERbeta-binding DNA regions*. Proc Natl Acad Sci U S A, 2008. **105**(7): p. 2604-9.
25. Stossi, F., et al., *Transcriptional profiling of estrogen-regulated gene expression via estrogen receptor (ER) alpha or ERbeta in human osteosarcoma cells: distinct and common target genes for these receptors*. Endocrinology, 2004. **145**(7): p. 3473-86.
26. Bai, Y. and V. Giguere, *Isoform-selective interactions between estrogen receptors and steroid receptor coactivators promoted by estradiol and ErbB-2 signaling in living cells*. Mol Endocrinol, 2003. **17**(4): p. 589-99.
27. Ali, S. and R.C. Coombes, *Endocrine-responsive breast cancer and strategies for combating resistance*. Nat Rev Cancer, 2002. **2**(2): p. 101-12.
28. Manas, E.S., et al., *Structure-based design of estrogen receptor-beta selective ligands*. J Am Chem Soc, 2004. **126**(46): p. 15106-19.
29. Barnett, D.H., et al., *Estrogen receptor regulation of carbonic anhydrase XII through a distal enhancer in breast cancer*. Cancer Res, 2008. **68**(9): p. 3505-15.
30. Choi, I., et al., *Human estrogen receptor beta-specific monoclonal antibodies: characterization and use in studies of estrogen receptor beta protein expression in reproductive tissues*. Mol Cell Endocrinol, 2001. **181**(1-2): p. 139-50.
31. Cowley, S.M., et al., *Estrogen receptors alpha and beta form heterodimers on DNA*. J Biol Chem, 1997. **272**(32): p. 19858-62.
32. Pettersson, K., et al., *Mouse estrogen receptor beta forms estrogen response element-binding heterodimers with estrogen receptor alpha*. Mol Endocrinol, 1997. **11**(10): p. 1486-96.
33. Carroll, J.S., et al., *Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1*. Cell, 2005. **122**(1): p. 33-43.
34. Kian Tee, M., et al., *Estradiol and selective estrogen receptor modulators differentially regulate target genes with estrogen receptors alpha and beta*. Mol Biol Cell, 2004. **15**(3): p. 1262-72.

## **CHAPTER 4**

### **A NETWORK OF ESTROGEN RECEPTORS $\alpha$ AND $\beta$ , AND COREGULATORS SRC3 AND RIP140, IN BREAST CANCER TRANSCRIPTIONAL REGULATION**

#### **4.1 ABSTRACT**

The nuclear hormone receptors, ER $\alpha$  and ER $\beta$ , are known to regulate the transcriptional response programs of their target cells, including breast cancer cells. However, their comparative abilities to localize at chromatin binding sites across the genome, and the association of these receptor binding sites with binding sites for the major coregulators SRC3 and RIP140, and with other transcription factors and chromatin remodeling factors is incompletely understood. Therefore, in these studies, we have used ChIP-seq in breast cancer cells containing three different complements of ERs (ER $\alpha$  alone, ER $\beta$  alone, or ER $\alpha$  + ER $\beta$ ) treated with estradiol to define the cartography of chromatin binding sites for ER $\alpha$ , ER $\beta$ , and the coregulators SRC3 and RIP140. We have found that ER $\alpha$  and ER $\beta$  bind to a similar, large number of sites (ca. 38,000) in cells containing only one ER subtype, with ER $\alpha$  sites being preferentially enriched in GATA and FOXA1 motifs, and ER $\beta$  sites preferentially enriched in E2F motifs. Although the two ERs occupy fewer sites in ER $\alpha$  + ER $\beta$  cells (ca. 20,000), the enrichment of their preferentially co-associated transcription factor motifs is further accentuated in the ER $\alpha$  and ER $\beta$  binding sites in these cells. Further, ER $\alpha$  was found, in general, to occupy chromatin sites more depleted of intrinsic nucleosomes compared to ER $\beta$ . Gene chip microarray transcriptional profiling and gene ontology analysis delineated a core set of genes that correlate with ER $\alpha$  proliferative and ER $\beta$  anti-proliferative effects in breast cancer cells. ER $\beta$  activation by estradiol was associated with the inhibition of genes associated with cell proliferation and the up-regulation of pro-apoptotic genes and genes responding to DNA damage, whereas ER $\alpha$  activation was associated with the downregulation of pro-apoptotic genes and genes repressing transcription. Analysis of chromatin

binding of SRC3 and RIP140 by ChIP-seq revealed that these coregulators are recruited preferentially to ER binding sites of estrogen-induced genes, whereas they are seldom recruited to ER binding sites of hormone-repressed genes, indicating that the SRC3-RIP140 complex is likely to be playing a central role in the induction of ER targeted genes. Our findings suggest an integrated model in which the actions of cofactors such as FOXA1, GATA3, and E2F enforce the selectivity and range of ER $\alpha$  and ER $\beta$  binding and gene regulatory actions, with the coregulators SRC3 and RIP140 preferentially supporting the stimulatory actions of both receptors on gene expression.

## 4.2 INTRODUCTION

Estrogens play pivotal roles in reproductive physiology, development and metabolism. Estrogens are also involved in several disease states, including breast and endometrical cancers as well as osteoporosis [1, 2]. Estrogen receptor alpha (ER $\alpha$ ) and beta (ER $\beta$ ) belong to the superfamily of nuclear transcription factors that mediate the actions of estrogens. It has been well-documented by various studies [3, 4] that the presence of ER $\alpha$  in breast cancer cells results in enhanced proliferation in response to estrogens. The role of ER $\beta$  and the manner in which it can impact estrogen mitogenicity in breast cancer, however, are less clear, although several reports have implicated ER $\beta$  as a negative regulator of ER $\alpha$  and shown that ER $\beta$  has antiproliferative effects in breast cancer cells [5-10]. As yet, however, no studies have convincingly identified genes that can explain ER $\alpha$  proliferative and ER $\beta$  anti-proliferative actions in breast cancer cells.

When ER $\alpha$  and ER $\beta$  bind estrogen, they undergo a conformation change that releases heat shock proteins and enhances ER dimerization and subsequent binding to the regulatory regions

of their target genes. Once bound on the chromatin, the activated ERs will recruit coregulators to form a multiprotein complex that activates the general transcriptional machinery which then regulates the expression of target genes [11, 12]. We have earlier demonstrated that the coregulators steroid receptor coactivator 3 (SRC3) and receptor interacting protein of 140kDa (RIP140) are recruited by both ER $\alpha$  and ER $\beta$  [4].

RIP140 is an atypical coregulator as it can act either as a coactivator or corepressor. Since its identification as a ER $\alpha$  coregulator [13], RIP140 has been found to interact with many other nuclear receptors, such as AR, GR, VDR, etc. [14-16]. RIP140 transactivation on ER $\alpha$  was observed in yeast [17, 18] and transient transfection of RIP140 expression plasmid leads to increase in ER $\alpha$  activity [13, 19]. Paradoxically, published data also indicate the repressive effects that RIP140 has on gene transcription. In one study, transrepression activity was observed on a reporter gene by full-length RIP140 fused to a GAL4-DBD [20], while global gene expression analysis in another study revealed that a significant number of genes were increased in RIP140-null cells as compared to RIP140-expressing wild-type cells [21]. SRC3, also known as amplified in breast cancer-1 (AIB1) [22], is a coactivator that promotes the transcriptional activity of ER $\alpha$  [22, 23]. Depletion of SRC3 led to a reduction of estrogen-stimulated proliferation in MCF-7 cells [24, 25].

Given the intimate relationships between ERs, coregulators, and gene expression, we hypothesize that we can better understand and dissect the differences in ER $\alpha$ - and ER $\beta$ -mediated transcriptional programs in breast cancer by an integrative genomic approach utilizing ER $\alpha$  and ER $\beta$  chromatin localization and the co-recruitment of SRC3 and RIP140 to ER binding sites. First, we employed chromatin immunoprecipitation (ChIP) combined with sequencing (ChIP-seq) to map the genomic landscape of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 in breast cancer cells

containing various complements of ER $\alpha$  and ER $\beta$ , in the presence of 17 $\beta$ -estradiol (E2). Next, gene expression microarray analyses were carried out to investigate the gene regulatory effects of ER $\alpha$  and ER $\beta$  in breast cancer cells. Finally, by correlating the global cartographies of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 with the results of the gene expression microarray analyses, we present evidence of ERs direct target genes that may contribute to ER $\alpha$  proliferative and ER $\beta$  anti-proliferative nature of these receptors in breast cancer cells. Our integrative ChIP-seq and expression profiling study extends beyond previous large-scale ER binding and gene expression studies [3, 4, 7-10, 26-32], because in this study we have comprehensively studied the recruitment of ERs (both ER $\alpha$  and ER $\beta$ , alone or together), together with their coregulators, to chromatin and the subsequent transcription impact effected by ERs on their target genes in breast cancer cells containing various complements of ER $\alpha$  and ER $\beta$ . We demonstrate that the selectivity and range of ER $\alpha$  and ER $\beta$  binding to the chromatin appear to depend on different cofactors. Our findings highlight the critical role of SRC3-RIP140 complex in breast cancer cell proliferation: The ER $\alpha$ -SRC3-RIP140 transcriptional program defines breast cancer proliferation, whereas the ER $\beta$ -SRC3-RIP140 program acts to positively regulate apoptosis in breast cancer cells.

#### 4.3 MATERIALS AND METHODS

##### *Ligands, Cell Culture and Adenovirus Infection*

MCF-7 cells were cultured in MEM (Sigma, St Louis, MO), supplemented with 5% calf serum (HyClone, Logan, UT), and 100  $\mu$ g/ml penicillin/streptomycin (Invitrogen, Carlsbad, CA). For estrogen-free experiments, the cells were maintained in phenol red-free MEM plus 5% charcoal-dextran-treated calf serum for at least 3 days and were then seeded at a density of 3 x

10<sup>5</sup> cells per 10 cm tissue culture dish (Corning, Corning, NY) for 2 days before adenovirus infection. Recombinant adenoviruses were constructed and prepared as described [10]. Conditions used were those described previously [10, 28, 33] to generate MCF-7 cells expressing levels of ER $\beta$  equal to that of the endogenously expressed ER $\alpha$ . Estradiol was from Sigma. The ER subtype-selective ligands, PPT and ERB-041, were synthesized as described [34, 35]. Studies used 10 nM E<sub>2</sub>, 50 nM PPT, and 500 nM ERB-041, minimal concentrations found to be maximally effective in binding to ER $\alpha$  and ER $\beta$  (E<sub>2</sub>), ER $\alpha$  (PPT), and ER $\beta$  (ERB-041), respectively, based on their relative binding affinities for receptor and their ability to regulate gene expression, based on our prior work [28].

#### *siRNA Transfection*

siRNA experiments were performed as previously described, and resulted in knockdown of ER $\alpha$  mRNA and protein by greater than 95% [4]. siER $\alpha$  sequences (Dharmacon) were: forward, 5'-UCAUCGCAUCCUUGCAAAdTdT-3', and reverse, 5'-UUUGCAAGGAAUGCGAUGAdTdT-3'.

#### *ChIP Assays*

ChIP for ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 were carried out as described [36] and used the following antibodies: ER $\alpha$  antibody HC-20 (Santa Cruz Biotechnology); ER $\beta$  antibodies were a combination with equal parts of CWK-F12 (produced by our lab) [37], GTX70182 (GeneTex), GR40 (Calbiochem), and PA1-311 (Affinity Bioreagents); SRC3 antibody H-270 (Santa Cruz Biotechnology); RIP140 antibody (Santa Cruz Biotechnology). The ChIP DNA was used for ChIP-seq analysis and quantitative real-time PCR.

#### *ChIP-seq Analyses*

The ChIP DNA was prepared into libraries and sequenced using the Genomic Analyzer following Illumina protocols. Sequences generated were mapped uniquely onto the human genome (hg18) by ELAND. MACS [38] was used to identify enriched peak regions with a p-value cutoff of  $6.0 \times 10^{-7}$  and FDR of 0.01.

#### *GeneChip mRNA transcriptional profiling microarrays*

Total RNA was used to generate cRNA, which was labeled with biotin according to techniques recommended by Affymetrix. The biotin-labeled cRNA was then hybridized to Affymetrix U133 plus 2.0 GeneChips, which contain oligonucleotide probe sets for over 47000 transcripts. After washing, the chips were scanned and analyzed using Affymetrix processing software. CEL files were processed using GeneSpring GX 11.0 software (Agilent) to obtain fold-change and p-value (with Benjamini and Hochberg multiple test correction) for each gene for each treatment relative to the vehicle control. We considered genes with fold-change  $> 1.5$  and p-value  $< 0.05$  as statistically significant, differentially expressed.

#### *Motif and GO category Analysis*

Overrepresented GO biological processes were determined by the web-based DAVID Bioinformatics Resources database [39, 40]. Motifs enrichment analysis was done using MotifEnrich program which uses TRANSFAC PWMs for motifs discovery (Ken Sung unpublished data).

## 4.4 RESULTS

### 4.4.1 Genes Are More Likely to be Up-Regulated in the Presence of ER $\beta$ as Compared to ER $\alpha$

We sought to determine ER $\alpha$  and ER $\beta$  mediated transcriptional responses in breast cancer cell by performing gene expression profiling. Cells were treated for 4 hours with estradiol



followed by hybridization of total RNA onto Affymetrix U133 plus 2.0 arrays (Figure 4.1A). For reference, since E2 activates both ERs, we also performed gene expression analysis using PPT, an ER $\alpha$ -selective ligand [35], and ERB-041, an ER $\beta$ -selective ligand [41] (Figure 4.1B). Triplicate experiments were performed, and analysis revealed that globally, genes were more likely to be up-regulated in the presence of ER $\beta$  compared to ER $\alpha$  (see Figure 4.1). For example, upon E2 treatment, 45% and 66% of the genes were up-regulated in cells containing ER $\alpha$  only ([ $\alpha$  cells]), or ER $\beta$  only ([ $\beta$  cells]) respectively. This pattern of response is further reinforced when we compared the gene expression profiles in cells containing both ER $\alpha$  and ER $\beta$  ([ $\alpha\beta$  cells]) under PPT and E2 treatment. Under conditions where ER $\beta$  is not activated (PPT treatment in [ $\alpha\beta$  cells]), we find the percentage of up- and down-regulated genes by PPT to be very similar to E2 treatment in [ $\alpha$  cells]. However, upon E2 treatment in [ $\alpha\beta$  cells] (which activated both ER $\alpha$  and ER $\beta$ ), we observe a reversal with more genes being up-regulated, indicating the preference for ER $\beta$  to up-regulate genes. In [ $\beta$  cells], we observed that globally ~65% of the genes were up-regulated by both E2 and ERB-041 treatments.

#### 4.4.2 Global Analysis of Cellular Processes Regulated Uniquely by ER $\alpha$ and ER $\beta$

Because ER $\beta$  has been implicated in the inhibition of cell proliferation while ER $\alpha$  is linked to proliferative activities [3, 7, 8], we therefore asked whether genes differentially regulated by ER $\alpha$  and ER $\beta$  have different underlying biological functions that can provide clues to the differences between the functional activities of the two ERs. Specifically, we focused our examination by looking at genes that are uniquely regulated by the respective ERs (see Figure 4.2A for the gene clusters that were examined) upon E2 treatment. We used DAVID [39, 40] to test for overrepresented gene ontology (GO) biological processes in the three gene clusters. The ER $\alpha$ -only gene cluster showed GO terms of “apoptosis” and “transcription repressor activity” as

overrepresented. Interestingly, most of the genes in these two GO categories are E2 repressed (see Figure 4.2B). By looking further into the GO descriptions of the group of apoptotic genes, we found 34% of these genes are pro-apoptotic, and 7% are anti-apoptotic (55% of the genes do not have further GO descriptions indicating whether they are pro- or anti-apoptotic). Our findings extend the results of previous studies [3, 10] by implicating that ER $\alpha$  proliferative activities may be through the down-regulation of apoptotic (mostly pro-apoptotic) and transcription repressor genes.

We found the GO terms of “programmed cell death” and “response to DNA damage” to be overrepresented in the ER $\beta$ -only cluster. Similarly, by looking further into the GO descriptions of the group of “response to DNA damage” genes, we found 33% are pro-cell death, and 12% are pro-proliferation (56% of the genes do not have further GO descriptions to indicate whether or not they are pro-proliferative or cell death inducing). The majority of the “programmed cell death” genes are pro-apoptotic (55% vs. 10% that are anti-apoptotic). In contrast to the ER $\alpha$ -only cluster, most of the ER $\beta$ -only cluster genes (“programmed cell death” and “response to DNA damage”) are up-regulated by E2 (see Figure 4.2C). The GO analysis suggest that ER $\beta$  activation may play a role in the inhibition of cell proliferation through the up-regulation of pro-apoptotic genes and genes responding to DNA damage.

#### 4.4.3 ER $\beta$ has the Ability to Modulate and to Replace ER $\alpha$ Functions

Motivated by the results in Section 4.4.2 and studies from other labs [3, 5-7, 9, 42] that show ER $\alpha$  proliferative and ER $\beta$  anti-proliferative actions in breast cancer cells, we selected several well-known cell-cycle and proliferation genes and compared their gene expression in cells containing various complements of ER $\alpha$  and ER $\beta$ . Cell cycle genes have been implicated directly in tumorigenesis, and their expression has been reported to coincide with cellular transformation

and oncogenic potential [43, 44]. The cell cycle and tumor proliferation genes (shown in Table 4.1) were selected from [45, 46].

We took the list of genes and extracted the patterns of gene expression from our E2-treated gene expression profile study (described in section 4.4.1). Interestingly, we observed that, similar to ER $\alpha$ , ER $\beta$  was able to induce a set of cell cycle and proliferation genes. However, this is not a complete replacement, as ER $\beta$  induction of these genes is mostly weaker than ER $\alpha$  induction (compare Figure 4.3A, ER $\alpha$  cells vs. ER $\beta$  cells). Taken together, our in vitro data suggests that ER $\beta$  is taking over some of the roles of ER $\alpha$  (upon the loss of ER $\alpha$  expression) in order to maintain normal MCF-7 cell functions. Also, ER $\beta$  is not inducing these genes as robustly as ER $\alpha$ , further indicating that ER $\beta$  is acting as, possibly, only a partial replacement. By contrast, while some genes are up-regulated by both ER $\alpha$  and ER $\beta$  (Figure 4.3B), most of the cell cycle and proliferation genes are down-modulated by ER $\beta$  (compare Figure 4.3A, ER $\alpha$  cells vs. ER $\alpha\beta$  cells) in the presence of ER $\alpha$ . Taken together, these results further reinforce our previous findings [4, 10] that ER $\beta$  has the ability to modulate ER $\alpha$ -mediated transcriptional activity and ER $\beta$  may attenuate MCF-7 growth by repressing ER $\alpha$ -regulated cell cycle and proliferative genes when both ERs are present together in the cells.

ER $\beta$  mainly plays an anti-proliferative role in breast cancer but apparently primarily in the presence of ER $\alpha$  where it modulates ER $\alpha$  signaling and represses ER $\alpha$  proliferative effects. However, upon the loss of ER $\alpha$ , ER $\beta$  might have the ability to replace some of ER $\alpha$  signaling in order to maintain normal cell functions. We therefore speculate that in MCF-7 cells, in addition to the ability of ER $\beta$  to modulate ER $\alpha$  signaling when both of them are present in the cells, ER $\beta$  can take over some of ER $\alpha$  signaling upon loss of ER $\alpha$  (i. e., in ER $\beta$ -only cells). This would indicate an extraordinary role of ER $\beta$  as a bifunctional switch that may be activating in one

setting and repressive in another. More detailed studies are required to elucidate the ability of ER $\beta$  to replace ER $\alpha$  functions upon the loss of ER $\alpha$  expression.

#### 4.4.4 Genome-Wide Analysis of ER $\alpha$ and ER $\beta$ Chromatin Binding by ChIP-Seq

In a preliminary study, we investigated ER-binding site selection by ER $\alpha$  and ER $\beta$  in human breast cancer cells [28] by ChIP-chip (Chapter 3). However, this initial assessment was limited by the fact that the analysis was not genome-wide, as the NimbleGen array was custom-designed based on known ER $\alpha$  binding sites and other computationally predicted sites. Therefore, the NimbleGen array used in our ChIP-chip study was not able to detect binding sites unique to ER $\beta$ . To obtain a more global and unbiased picture of ER $\alpha$  and ER $\beta$  binding sites, we generated ChIP-seq libraries to study the genome-wide localization of ER $\alpha$  and ER $\beta$ , when present separately or together, in response to E2 treatment (Methods, Table 4.2). Due to the unbiased and genome-wide manner of sequencing, the ChIP-seq ER datasets described in this chapter should represent all ER $\alpha$  and ER $\beta$  binding sites in MCF-7 cells. Therefore, we wished to validate the two major findings from our ChIP-chip study, namely (i) the constriction of ER $\alpha$  binding sites by ER $\beta$  and (ii) the shunting of ER $\beta$  to less “favorable” ERE binding sites, using the more comprehensive ChIP-seq datasets.

First, we observed from the ChIP-chip study that in cells containing both ER $\alpha$  and ER $\beta$ , the presence of ER $\beta$  restricts the range of binding sites for ER $\alpha$  when compared to that seen in ER $\alpha$ -only cells (3252 ER $\alpha$  [ $\alpha\beta$  cells] vs. 4405 ER $\alpha$  [ $\alpha$  cells] ChIP-chip sites). In contrast, the number of ER $\beta$  sites assessed by ChIP-chip are not constricted by ER $\alpha$  (1744 ER $\beta$  [ $\alpha\beta$  cells] vs. 1897 ER $\beta$  [ $\beta$  cells] ChIP-chip sites). Comparing with the ChIP-seq data (see Table 4.2), we observe the restriction of ER $\alpha$  binding sites by ER $\beta$  (19950 ER $\alpha$  [ $\alpha\beta$  cells] vs. 37898 ER $\alpha$  [ $\alpha$  cells] ChIP-seq sites), which is in agreement with the ChIP-chip study. However, ER $\beta$  binding sites are also

restricted by ER $\alpha$  in our ChIP-seq study (20288 ER $\beta$  [ $\alpha\beta$  cells] vs. 37757 ER $\beta$  [ $\beta$  cells] ChIP-seq sites). The restriction of ER $\beta$  was not seen in our ChIP-chip study. It is possible that the difference between the two studies is due to a lack of a number of ER $\beta$  binding sites that were detected by the custom-designed NimbleGen array (ChIP-chip study). As the NimbleGen array was designed using known ER $\alpha$  sites, binding sites unique to ER $\beta$  would not have been detected, whereas ChIP-seq is able to detect most of the ER $\beta$  binding sites. Therefore, the ChIP-chip study might not present a complete picture of ER $\beta$  binding in the genome, leading us to not detect the constriction of ER $\beta$  binding sites by ER $\alpha$ .

Second, the ChIP-chip study indicated that ER $\alpha$  and ER $\beta$  are in competition for binding sites, with ER $\alpha$  “driving” ER $\beta$  to occupy sites less enriched in ER motif. Similarly, we also observed the dominant nature of ER $\alpha$  in our ChIP-seq study. In cells containing both ERs, we found that only 9% of sites unique to ER $\beta$  have an ER motif (Table 4.3D). This is in contrast to other ER sites which generally have 14-24% enrichment of ER motif (Table 4.3A, 4.3B, 4.3C, and 4.3E). Therefore, the shunting of ER $\beta$  by ER $\alpha$  to sites less enriched in ER motifs is also observed in the ChIP-seq data.

As shown, despite the differences in technology, the ER binding site datasets generated by ChIP-seq are in good agreement with the earlier ChIP-chip data, because we are able to observe both the restriction of ER sites and the occupancy of sites less enriched in ER motifs by ER $\beta$  in the ChIP-seq data.

#### 4.4.5 ER $\alpha$ , But Not ER $\beta$ , has Enriched GATA and FOXA1 Motifs

To determine whether there are any differences in the genomic sequences to which ER $\alpha$  and ER $\beta$  bind, we scanned the ERs sites for enriched motifs using MotifEnrich (Ken Sung unpublished data). Not surprisingly, our search revealed the ER response element as the topmost

overrepresented motif and identified several other motifs that had been found in previous studies [28, 31], such as BACH2, AP-1, SP-1 etc (Table 4.4 to Table 4.7). However, a comparison of the enriched motifs found near ER $\alpha$  [ $\alpha$  cells] and ER $\beta$  [ $\beta$  cells] binding sites revealed a surprising result – both FOXA1 and GATA were overrepresented in ER $\alpha$  sites but not in ER $\beta$  sites (see Table 4.3). This is rather unexpected, as studies [26, 27, 47, 48] have shown that FOXA1 and GATA3 are E2-regulated genes, and both factors are crucial for ER $\alpha$  chromatin binding and subsequent activation of its target genes. In order to address the issue that some of the sites bound by GATA3 and FOXA1 might not contain conserved motifs recognizable by TRANSFAC PWMs, we analyzed the overlap of ER $\alpha$  [ $\alpha$  cells] and ER $\beta$  [ $\beta$  cells] binding sites with the recently published FOXA1 ChIP-chip data [47] and the ChIP-seq data for FOXA1 and GATA3 (GIS unpublished data, SayLi KONG, Guoliang Li and Edison Liu). We were therefore expecting ER $\beta$  to utilize either FOXA1 or GATA3 as cofactors. We found that both FOXA1 (p-value < 1e-257) and GATA3 (p-value < 5e-313) binding sites have a higher propensity to overlap ER $\alpha$  [ $\alpha$  cells] binding sites as compared to ER $\beta$  [ $\beta$  cells] binding sites (see Figure 4.4A-4.4C). For example, using the GIS FOXA1 and GATA3 data, we found that 20.4% of ER $\alpha$  binding sites and 11.3% of ER $\beta$  binding sites have FOXA1 binding in close proximity (Figure 4.4B) and 30.8% of ER $\alpha$  binding sites and 13.8% of ER $\beta$  binding sites have GATA3 binding in close proximity (Figure 4.4C). The result is even more striking if we separate ER $\alpha$  [ $\alpha$  cells] and ER $\beta$  [ $\beta$  cells] sites into ER $\alpha$ -only [ $\alpha$  cells], ER $\beta$ -only [ $\beta$  cells], and sites *in common* to both ERs (i.e., can be occupied by *either* ER $\alpha$  or ER $\beta$  when they are present alone) (see Figure 4.5A for the Venn separation diagram). Only ~3% of sites which are exclusively occupied by ER $\beta$  are co-occupied by either FOXA1 or GATA3. In contrast, 18.6% to 33.9% of the sites exclusively bound by ER $\alpha$  are co-occupied by either FOXA1 or GATA3. (see Figure 4.6A-4.5C).

We then turned our attention at ER binding sites in [ $\alpha\beta$  cells], and we separated ER $\alpha$  [ $\alpha\beta$  cells] and ER $\beta$  [ $\alpha\beta$  cells] binding sites into sites where ER $\alpha$  binds uniquely (ER $\alpha$ -only[ $\alpha\beta$  cells]), ER $\beta$  binds uniquely (ER $\beta$ -only[ $\alpha\beta$  cells]), and sites that can be shared by both ERs (ER $\alpha$  [ $\alpha\beta$  cells]&ER $\beta$ [ $\alpha\beta$  cells]) (see Figure 4.5B for the Venn separation diagram). We again observed that FOXA1 (p-value < 2e-70) and GATA3 (p-value < 2e-160) are recruited more frequently to ER $\alpha$  sites than to ER $\beta$  sites (see Figure 4.7A-4.7C). Overall, our genome-wide location analyses suggest the cofactors FOXA1 and GATA3 do not co-occupy ER $\beta$  binding sites as frequently as ER $\alpha$  binding sites.

#### 4.4.6 E2F Motif is Associated with ER $\beta$ Binding

Although the E2F motif was overrepresented both in the binding sequences of ER $\alpha$  [ $\alpha$  cells] and ER $\beta$  [ $\beta$  cells], we observed that E2F is three times more enriched in ER $\beta$  [ $\beta$  cells] compared to ER $\alpha$  [ $\alpha$  cells] binding sites (see Table 4.8A). By plotting the occurrences of the E2F motif around ER binding sites, we observed that the E2F motif is more likely to be located on ER $\beta$  [ $\beta$  cells] peaks as compared to ER $\alpha$  [ $\alpha$  cells] peaks (Figure 4.8A). We were therefore interested to know whether a similar distribution of E2F motif enrichment occurs across ERs binding sites in [ $\alpha\beta$  cells]. We again separated the binding sites into ER $\alpha$ -only[ $\alpha\beta$  cells], ER $\beta$ -only[ $\alpha\beta$  cells], and sites that can be shared by both ERs, and we analyzed these three ER binding site sets for enriched E2F motifs (see Table 4.8B and Figure 4.8B). Interestingly, we were not able to find any enrichment of E2F motif around the sites occupied by ER $\alpha$ -only[ $\alpha\beta$  cells], whereas a strong enrichment of E2F motif is observed in sites occupied by ER $\beta$ -only[ $\alpha\beta$  cells], suggesting that the E2F family of transcription factors might be recruited more frequently to ER $\beta$  sites. It is possible that our motif scanning did not reveal the full extent of E2F binding in the genome as (i)

E2F can binds to imperfect E2F motifs, (ii) the TRANSFAC motif for E2F is unable to predict all the members of E2F binding, or (iii) the conservative p-value cutoff ( $p\text{-value} < 1e-4$ ) from our motif scanning program limits the number of E2F sites called. Hence, despite our results here, it is still conceivable that the E2F family members can be recruited by both ERs.

The E2F family members consist of activators and repressors, with E2F1-E2F3a generally thought of as transcription activators and E2F3b-E2F8 as transcription repressors. It is possible that ER $\alpha$  and ER $\beta$  are recruiting different members of the E2F family, thus leading to the different transcription regulation programs that are observed between the two ERs in breast cancer cells. More detailed studies are required to distinguish the recruitment of members of E2F family by either ER. Nonetheless, the fact that our genome-wide analysis of ERs binding sites revealed enrichment of E2F motif around ER $\beta$  sites raises the possibility that E2F and ER $\beta$  might share a significant fraction of their *cis*-regulatory sites.

#### 4.4.7 ER $\beta$ is Occupying “Less Accessible” Sites

MotifEnrich scanning indicates that 14-24% of the ERs binding sites contain the ERE motif (see Table 4.3), but interestingly, only 9% of the sites at ER $\beta$ -only [ $\alpha\beta$  cells] have an ERE motif (see Table 4.3D). This is consistent with what we saw previously [28], where ER $\beta$  sites are found to be shunted to sites less enriched in ERE motifs in the presence of ER $\alpha$ . Nucleosome depletion has been reported around transcription factor binding sites [49, 50], and it has been suggested that nucleosome depletion assists in directing transcription factors to bind to appropriate regulatory sites. Considering our observations that DNA sequences at ER $\beta$  sites are less enriched in GATA3 and FOXA1 motifs together with the fact that ER $\beta$ -only [ $\alpha\beta$  cells] are shunted to sites less enriched in ERE motif, we hypothesized that ER $\beta$ , in general, might be occupying less favorable or less accessible chromatin sites as compared to ER $\alpha$ . Therefore, we



applied intrinsic nucleosome preference model proposed by Kaplan and colleagues [51] to estimate nucleosome occupancy at ER $\alpha$  and ER $\beta$  sites. For comparison, we also looked at the intrinsic nucleosome occupancy of FOXA1 sites defined by ChIP-chip study [47]. Kaplan and colleagues first measured the genome-wide occupancy of nucleosomes assembled on purified yeast genomic DNA. Using this data, they devised a computational model, the intrinsic nucleosome preference model, which predicts nucleosome occupancy solely based on the DNA sequence. Therefore, by using the nucleosome preference model as proposed by Kaplan, we should be able predict the intrinsic nucleosome occupancy across ER $\alpha$  and ER $\beta$  binding sites. Using this model, we found that ER $\alpha$  sites (both in [ $\alpha$  cells] and [ $\alpha\beta$  cells]) are more depleted of intrinsic nucleosomes as compared to ER $\beta$  sites (in [ $\beta$  cells] and [ $\alpha\beta$  cells]) (see Figure 4.9). This indicates that, in MCF-7 cells, ER $\alpha$  is binding to more accessible DNA sequences (ones with fewer intrinsic nucleosomes) as compared to ER $\beta$ .

#### 4.4.8 The Hierarchy of ERs Chromatin Binding Sites

Upon binding to chromatin, ER $\alpha$  and ER $\beta$  will recruit coregulators to form a complex that is competent to effect transcription of the target genes. To characterize the transcriptional pathways among ERs and its coregulators, we assessed the chromatin recruitment of two of ER's coregulators, SRC3 and RIP140, in MCF-7 cells that contain ER $\alpha$  separately, ER $\beta$  separately or both ERs, in response to E2 treatment (Methods, Table 4.9). SRC3, which is a member of the p160 family, is a classic coactivator of several nuclear hormone receptors including ERs [22]. RIP140 [13] is another nuclear hormone coregulator and has been shown to interact with ER $\alpha$  and modulate estrogen-induced gene activity [52, 53]. Since SRC3 and RIP140 are non-DNA binding coregulators that are recruited by ERs, we therefore expected that SRC and RIP140 recruitment to chromatin would be dependent on ERs. Figure 4.10 shows that within each of the

MCF-7 sub-lines, most of the SRC3 and RIP140 sites are co-localized with ER binding sites.

There does not seem to be a bias in recruiting SRC3 or RIP140 by either of the ERs.

As there are more ER binding sites in the genome as compared to E2 regulated genes, we hypothesized that there is a hierarchy to ER chromatin binding sites, ranging from sites that are required for gene regulation to redundant sites. ER-mediated transcription can be thought of as a two-step process. First, upon activation by ligands, ERs searches for and binds to the regulatory regions of their target genes. Then, after binding to DNA, coregulators are recruited by the ERs to carry out all the subsequent biochemical reactions required for induction or repression of the genes. Because the transcriptional potency of the ERs depends on coregulator recruitment, ER genomic binding sites where both ER and coregulators are localized should represent the “most functional” sites of the ER chromatin binding sites. Identifying these sites should enable us to pinpoint important genes occupying the driver seats of ERs actions. We therefore sought to identify ER target genes through the overlap of E2-regulated genes with the genomic localization of the ERs and the coregulators SRC3 and RIP140. This should allow us to identify a set of high confidence ER direct target genes which we can use to understand and elucidate the gene regulatory functions of the ERs and the mechanisms behind the opposing roles that appear to be played by ER $\alpha$  and ER $\beta$  in breast cancer.

#### 4.4.9 SRC3 and RIP140 are Recruited to E2-Induced Genes by ERs

We first segregated ERs sites into four categories (see Table 4.10): ER that co-localized on chromatin with both SRC3 and RIP140 (E+S+R), ER that co-localized on chromatin with either SRC3 or RIP140 (E+S and E+R, respectively), and ER sites that do not have co-localized SRC3 or RIP140 (E). We then proceeded to define ER target genes using the different ER binding sites categories. Specifically, we classified ERs target genes as E2-regulated genes whose regulatory

regions (20kbp up- and down-stream of TSS) are bound by ERs and/or coregulators (sites from E+S+R, E+S, E+R, and E). The result (see Table 4.11) is especially striking as there is a strong bias for genes that are occupied by ER and its coregulators (genes bound to E+S+R, E+S, and E+R sites) to be E2 induced. The correlation between ER and coregulators occupancy with gene induction was the most pronounced at sites where ERs are bound together with SRC3 and RIP140 (genes with E+S+R sites). Collectively, our results suggest that ERs are quite capable of recruiting SRC3 and RIP140 to E2-induced genes, whereas SRC3 and RIP140 are seldom recruited by ERs to E2 repressed genes.

SRC3 has been shown to promote the transcriptional activity of ER $\alpha$  [22, 23], and its depletion leads to the reduction of estrogen-stimulated proliferation of MCF-7 cells [24, 25]. RIP140 appears to be an atypical coregulator, as published data indicates that it can act either as a coactivator [13, 17, 19] or corepressor [20, 21]. It is therefore interesting to see that on a genome-wide scale, RIP140, despite its potential to act as a corepressor, is acting in concert with the coactivator SRC3 to positively regulate ER-mediated gene transcription in breast cancer cells.

#### 4.4.10 The ERs Together with SRC3 and RIP140 Define the ER-Mediated Transcriptional Program

Several studies done by us and others have used breast cancer cells to examine ER $\alpha$ - and ER $\beta$ -mediated gene expression programs [3, 4, 8-10, 29]. To date, a myriad of biological processes ranging from cell-cell signaling, ion homeostasis, proliferation, control of cell cycle to apoptosis have been associated with genes regulated by ERs. However, no studies have heretofore convincingly identified genes that can explain ER $\alpha$  proliferative and ER $\beta$  anti-proliferative actions in breast cancer cells. To test the hypothesis that genes identified by us

(genes with E+S+R, E+S, and E+R binding sites, Table 4.11) represent important ER target genes, we interrogated these sets of genes for overrepresented gene ontology (GO) biological processes. We focused our analysis on genes from [ $\alpha$  cells] and [ $\beta$  cells] as these genes should be directly targeted by ER $\alpha$  and ER $\beta$ , respectively, as the other ER subtype is not present together in the cell. As both SRC3 and RIP140 appeared to be acting in a cooperative manner with the ERs in enhancing gene regulation, we therefore combined genes from categories E+S+R, E+S and E+R (Table 4.11) into a single set for our GO analysis. Among the genes that are associated with sites occupied by ER $\alpha$  and its coregulators in [ $\alpha$  cells], the GO terms “**Regulation of cell proliferation**” (GO 0042127) followed by “**Positive regulation of cell proliferation**” (GO 0008284) were the most overrepresented (Table 4.12A) GO terms. It is noteworthy that the most enriched GO biological function in the set of ER $\alpha$  direct target genes is cell proliferation. This suggests that the main ER $\alpha$ -mediated regulatory pathway in breast cancer cells may be through the positive regulation of cell proliferation.

In contrast, “**Regulation of apoptosis**” (GO 0042981), “**Regulation of programmed cell death**” (GO 0043067), “**Regulation of cell death**” (GO 0010941), “**Positive regulation of apoptosis**” (GO 0043065), “**Positive regulation of programmed cell death**” (GO 0043068), and “**Positive regulation of cell death**” (GO 0010942) dominated the GO list for genes that are associated with sites occupied by ER $\beta$  and its coregulators in [ $\beta$  cells] (Table 4.12B). The identification of apoptosis as the most enriched GO biological term in ER $\beta$  target genes suggests that the phenotypic consequences of ER $\beta$ -mediated transcriptional activity are in direct competition with those of ER $\alpha$ . The results presented here are complementary to those of our earlier GO analysis using genes uniquely regulated by ERs (section 4.4.2). Based on our combinatorial expression and binding sites analysis, we put forth the hypothesis that the

differences between the ER $\alpha$ - and ER $\beta$ -mediated activities observed in breast cancer cell are the consequence of the opposing groups of genes regulated by ER $\alpha$  and ER $\beta$ : ER $\alpha$  is inducing cell proliferative genes to effect its tumorigenic activity, whereas ER $\beta$  anti-proliferative activity is obtained by the up regulation of apoptotic genes. Our earlier analysis identified the down-regulation of pro-apoptotic genes by ER $\alpha$  and the up-regulation of pro-apoptotic genes by ER $\beta$ , changes that enforce our conclusion here.

#### 4.5 DISCUSSION

In this study, we report the genome-wide identification of ER $\alpha$ , ER $\beta$ , and coregulators SRC3, and RIP140 chromatin localization in multiple sub-lines of MCF-7 breast cancer cells, expressing various complements of ER $\alpha$  and ER $\beta$ , in the presence of E2. In addition, we carried out microarray expression profiling experiments to investigate the gene regulatory effects of ER $\alpha$  and ER $\beta$  in breast cancer cells in the presence of E2 and ER-subtype selective ligands. We generated a compendium of 10 genome-wide ChIP-seq datasets and a set of paired gene expression data which we used to dissect the complex transcriptional programs orchestrated by both ER $\alpha$  and ER $\beta$  in breast cancer cells. By analyzing the genome-wide localization data of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 with global gene expression, we were able to make several novel observations.

First, we showed that FOXA1 and GATA3 are not as frequently co-occupied at ER $\beta$  chromatin binding sites compared to ER $\alpha$  chromatin binding sites. In contrast, we found the E2F motif to be highly enriched in ER $\beta$  but not ER $\alpha$  binding sites. Both FOXA1 and GATA3 are important DNA-bound cofactors of ER $\alpha$  [26, 27, 47, 48], and they have been shown to be necessary for ER $\alpha$  chromatin binding and subsequent activation of its target genes. The E2F

family members consist of activators and repressors, and the fact that our results show that ER $\beta$  binding sites are more enriched with E2F motifs raises the possibility that ER $\beta$  might be preferentially binding to the chromatin near E2F binding sites and utilizing some of the E2F members to affect the regulation of its target genes. In this regard, despite the fact that both ER $\alpha$  and ER $\beta$  have similar DNA-binding domains and thus can recognize the same estrogen response element motif on the DNA [54-57], our data suggests that there are other cofactors such as FOXA1, GATA3, E2F, etc, enforcing the selectivity and range of ER $\alpha$  and ER $\beta$  binding.

Second, using the intrinsic nucleosome preference model proposed by Kaplan et al. [51], we showed that ER $\alpha$  is in general occupying chromatin sites that are more depleted of intrinsic nucleosomes compared to ER $\beta$  sites. This might, in part, explain why ER $\alpha$  has a more dominant presence in breast cancer cells, compared to ER $\beta$ , as it able to bind more easily to chromatin to regulate its target genes. This observation fits well with our previous study [28], where we proposed that ER $\alpha$  may be dominating ER $\beta$  in terms of binding to DNA.

Third, by co-localizing ERs, SRC3 and RIP140 binding to gene expression, we found a strong bias for the recruitment of SRC3 and RIP140 by both ER $\alpha$  and ER $\beta$  to the regulatory regions of E2-induced genes. SRC3 recruitment by ER $\alpha$  to E2-induced genes was reported previously by others [30, 58]. It is interesting the RIP140, which can functions as either a coactivator or a corepressor, is acting in concert with SRC3 to induce gene transcription. Lanz et al. [58] has previously reported the association of corepressor proteins such as NURD, HDACs with SRC3 in E2-stimulated MCF-7 wild type cells. Studies have also shown that it is possible for corepressors to enhance transcription [59, 60]. It is therefore conceivable that a new unidentified role for SRC3 is its ability to activate transcription through the recruitment of corepressors in breast cancer cells. We also note that there is hardly any recruitment of SRC3 or

RIP140 to E2-repressed genes, which suggests that SRC3 and RIP140 are seldomly used by ER for gene repression in breast cancer cells. Our study has provided new insights into the underlying mechanism of ER-mediated transcriptional program, namely, that the SRC3-RIP140 complex might be playing a central role in the induction of ER targeted genes in breast cancer cells.

Finally, the co-localization of ERs, SRC3 and RIP140 binding to gene expression has allowed us to identify a set ER $\alpha$  and ER $\beta$  target genes. Various gene expression studies were done in an attempt to identify genes which can explain the divergent roles of ER $\alpha$  and ER $\beta$  in breast cancers. As yet, no studies have convincingly identified genes that can explain ER $\alpha$  proliferative and ER $\beta$  anti-proliferative actions, because all of the studies used the whole set of E2-stimulated genes and not the subset of ER-targeted genes. However, with our genome-wide location and expression analyses, we are able to filter the E2-stimulated genes and identify a core set of ER target genes. GO analysis showed that genes occupied by ER $\alpha$ , SRC3, and/or RIP140 in [ $\alpha$  cells] have the GO term of pro-proliferation as the most statistically enriched. Similarly, genes occupied by ER $\beta$ , SRC3, and/or RIP140 in [ $\beta$  cells] are statistically enriched with the GO term of pro-apoptotic. These sets of genes may, at least in part, be responsible for the distinct roles of ER $\alpha$  and ER $\beta$  in the regulation of breast cancer cell proliferation.

In summary, our study has defined a complete list of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 chromatin interaction sites in MCF-7 breast cancer cells and has delineated a network of interactions among these regulatory factors. Given the fact that the predominant hormone in breast cancer biology is E2, it is therefore important for us to understand the nature of its receptors - ER $\alpha$  and ER $\beta$  - in terms of their chromatin binding, their coregulator recruitment, and the subsequent regulation of ER-target genes. Coupled with paired gene expression data, we

were able to identify a core set of genes that correlate with ER $\alpha$  proliferative and ER $\beta$  anti-proliferative effects. Our findings provide working models in which ER $\alpha$ -SRC3-RIP140-mediated transcription plays a crucial role in breast cancer cell proliferation, whereas ER $\beta$ -SRC3-RIP140-mediated transcription induces apoptosis in breast cancer.



#### 4.6 TABLES AND FIGURES

**Table 4.1.** List of cell cycle genes taken from [46]. These cell cycle genes were then correlated with [45] to identify proliferation genes.

<b>Genes</b>	<b>Assigned cell cycle phase [46]</b>	<b>Identified as breast tumors proliferation genes in [45]</b>
E2F5	G2/M	
CCNE1	G1/S	Proliferation
CCNE2	G1/S	
CDC25A	G1/S	
MCM6	G1/S	Proliferation
SLBP	G1/S	
NASP	G1/S	Proliferation
TYMS	S phase	Proliferation
BIRC5	G2/M	Proliferation
CDC25B	G2/M	Proliferation
AURKA	G2/M	Proliferation
TOP2A	G2	Proliferation
RRM2	S phase	Proliferation
BUB1B	G2/M	
CCNB2	G2/M	
CENPA	G2/M	

**Table 4.2.** Summary of ER binding sites in the three MCF-7 cells expressing ER $\alpha$  only, both ER $\alpha$  and ER $\beta$ , or ER $\beta$  only.

	<b>[<math>\alpha</math> cells]</b>	<b>[<math>\alpha\beta</math> cells]</b>	<b>[<math>\beta</math> cells]</b>
Total ER $\alpha$ binding sites	37898	19950	NA
Total ER $\beta$ binding sites	NA	20228	37757

**Table 4.3.** Evaluation of ER, FOXA1 and GATA motifs and their enrichment at ER binding sites.

**A.** Enrichment of ER, FOXA1, and GATA motifs at ER $\alpha$  [ $\alpha$  cells] binding sites

<b>TRANSFAC ID</b>	<b>% of ER<math>\alpha</math> [<math>\alpha</math> cells] with motif</b>	<b>p-value</b>
V_ER_Q6	14.31%	0
V_HNF3ALPHA_Q6	1.53%	4.91E-91
V_GATA1_05	2.19%	1.04E-139

**B.** Enrichment of ER, FOXA1, and GATA motifs at ER $\beta$  [ $\beta$  cells] binding sites

<b>TRANSFAC PWM</b>	<b>% of ER<math>\beta</math> [<math>\beta</math> cells] with motif</b>	<b>p-value</b>
V_ER_Q6_02	16.36%	0
V_HNF3ALPHA_Q6	0.00%	1.00E+00
V_GATA1_05	0.08%	2.36E-01

**C.** Enrichment of ER, FOXA1, and GATA motifs at ER $\alpha$ -only [ $\alpha\beta$  cells] binding sites

<b>TRANSFAC PWM</b>	<b>% of ER<math>\alpha</math>-only [<math>\alpha\beta</math> cells] with motif</b>	<b>p-value</b>
V_ER_Q6	23.94%	0
V_HNF3ALPHA_Q6	0.71%	2.37E-15
V_GATA1_05	1.17%	1.29E-25

**D.** Enrichment of ER, FOXA1, and GATA motifs at ER $\beta$ -only [ $\alpha\beta$  cells] binding sites

<b>TRANSFAC PWM</b>	<b>% of ER<math>\beta</math>-only [<math>\alpha\beta</math> cells] with motif</b>	<b>p-value</b>
V_ER_Q6_02	8.86%	0
V_HNF3ALPHA_Q6	0.00%	1.00E+00
V_GATA1_05	0.00%	1.00E+00

**E.** Enrichment of ER, FOXA1, and GATA motifs at ER $\alpha$  [ $\alpha\beta$  cells] & ER $\beta$ [ $\alpha\beta$  cells] binding sites

<b>TRANSFAC PWM</b>	<b>% of ER<math>\alpha</math> [<math>\alpha\beta</math> cells] &amp; ER<math>\beta</math> [<math>\alpha\beta</math> cells] with motif</b>	<b>p-value</b>
V_ER_Q6	22.28%	0
V_HNF3ALPHA_Q6	0.00%	1.00E+00
V_GATA1_05	0.28%	8.07E-05

**Table 4.4.** TFBS enrichment in ER $\alpha$ [ $\alpha$  cells] binding sites.

Rank	Motif Family	Transfac ID	% of sites with motif	p-value
1	ERE	V_ER_Q6	14.31%	0
2	BACH	V_BACH2_01	5.56%	0
3	AP1	V_AP1_01	5.54%	0
4	AP2	V_AP2ALPHA_02	4.61%	0
5	FOX	V_FREAC4_01	4.38%	0
6	CREB	V_ATF3_Q6	3.05%	0
7	NRF	V_NRF2_Q4	2.95%	0
8	AR	V_AR_01	2.88%	0
9	DBP	V_DBP_Q6	2.80%	0
10	GATA	V_GATA_Q6	2.51%	4.87E-152
11	MEF3	V_MEF3_B	2.31%	4.12E-157
12	NF1	V_NF1_Q6	1.98%	1.97E-105
13	LMAF	V_LMAF_Q2	1.97%	2.06E-123
14	VMAF	V_VMAF_01	1.88%	1.61E-110
15	FXR	V_PXR_Q2	1.70%	1.96E-76
16	LRH1	V_LRH1_Q5	1.69%	1.97E-68
17	PAX	V_PAX6_Q2	1.56%	3.37E-96
18	EBOX	V_USF_Q6_01	1.55%	2.71E-85
19	HIC1	V_HIC1_03	1.46%	3.77E-73
20	MYOGNF1	V_MYOGNF1_01	1.33%	1.43E-69
21	MAF	V_MAF_Q6_01	1.29%	1.43E-67
22	SMAD	V_SMAD4_Q6	1.22%	2.31E-65
23	SP1	V_SP1_Q6_01	1.20%	4.22E-51
24	CACCC	V_CACCCBINDINGFACTOR_Q6	1.07%	2.69E-47
25	MYB	V_MYB_Q6	1.06%	1.05E-47
26	ETS	V_NERF_Q2	1%	2.24E-45
27	GATA DIMER	V_EVI1_06	1%	4.37E-45
28	CP2	V_CP2_01	0.94%	6.06E-41
29	AP4	V_AP4_Q6_01	0.92%	2.47E-38
30	TEF	V_TEF1_Q6	0.91%	3.93E-13
31	MEIS1	V_MEIS1_01	0.90%	7.90E-35
32	CMAF	V_CMAF_01	0.90%	3.05E-26
33	ROAZ	V_ROAZ_01	0.85%	5.82E-39
34	P53	V_P53_02	0.78%	2.98E-36
35	MIF1	V_MIF1_01	0.75%	3.35E-37
36	HEN	V_LBP1_Q6	0.72%	1.19E-24
37	E2F	V_E2F1_Q4	0.70%	9.09E-35
38	TGTGGT	V_PEBP_Q6	0.68%	9.90E-27
39	KAISO	V_KAISO_01	0.67%	3.22E-28
40	OLF1	V_OLF1_01	0.64%	6.87E-27
41	TGIF	V_TGIF_01	0.55%	2.73E-18
42	STAT	V_STAT1_01	0.55%	1.62E-26
43	PTF1BETA	V_PTF1BETA_Q6	0.53%	5.79E-18

**Table 4.4. (cont.)**

<b>Rank</b>	<b>Motif Family</b>	<b>Transfac ID</b>	<b>% of sites with motif</b>	<b>p-value</b>
44	EGR	V_EGR1_01	0.48%	3.43E-20
45	SOX	V_SOX10_Q6	0.47%	3.22E-07
46	NFKB	V_NFKAPPAB65_01	0.47%	7.03E-18
47	RFX	V_RFX1_01	0.46%	5.11E-14
48	XPF1	V_XPF1_Q6	0.45%	2.14E-40
49	SP3	V_SP3_Q3	0.44%	1.44E-24
50	EBF	V_EBF_Q6	0.38%	5.99E-17
51	IK	V_IK3_01	0.37%	1.29E-13
52	SEF1	V_SEF1_C	0.37%	4.64E-17
53	DEAF1	V_DEAF1_01	0.35%	4.27E-22
54	XVENT1	V_XVENT1_01	0.34%	1.70E-12
55	P300	V_P300_01	0.32%	7.17E-18
56	CEBP	V_CEBP_01	0.31%	1.91E-06
57	CACCT	V_AREB6_01	0.30%	3.78E-15
58	ZF5	V_ZF5_B	0.29%	1.87E-21
59	LEF	V_LEF1_Q2_01	0.16%	0.000524
60	E2	V_E2_Q6	0.15%	3.78E-08
61	TATA	V_MTATA_B	0.13%	2.26E-06
62	MINI	V_MINI19_B	0.12%	4.42E-14
63	LDSPOLYA	V_LDSPOLYA_B	0.12%	9.41E-05
64	ALX4	V_ALX4_01	0.12%	1.68E-06
65	CAAT	V_YY1_02	0.11%	1.05E-06
66	WHN	V_WHN_B	0.10%	1.53E-12
67	SREB	V_SREBP_Q3	0.09%	0.000516
68	Initiator	V_GEN_INI3_B	0.07%	1.94E-07
69	AHR	V_AHR_01	0.06%	0.120647
70	IPF	V_IPF1_Q4	0.06%	0.000105
71	ZEC	V_ZEC_01	0.04%	0.124365
72	HOX	V_HOXA7_01	0.03%	0.000182
73	BRACH	V_TBX5_02	0.03%	0.044359
74	GCM	V_GCM_Q2	0.01%	7.30E-05
75	OCT	V_OCT1_06	0.01%	0.037178
76	GGG	V_CHCH_01	0%	1.21E-05

We used MotifEnrich to scan for enriched motifs in ER $\alpha$ [ $\alpha$  cells] binding sites. The motifs are ranked according to their enrichment (descending order) in ER $\alpha$ [ $\alpha$  cells] binding sites.

**Table 4.5.** TFBS enrichment in ER $\alpha$ [ $\alpha\beta$  cells] binding sites.

Rank	Motif Family	Transfac	% of sites with motif	p-value
1	ERE	V_ER_Q6	23.05%	0
2	CREB	V_ATF3_Q6	4.23%	0
3	AP1	V_AP1_Q6	4.17%	0
4	BACH	V_BACH2_01	4.08%	0
5	AR	V_AR_01	3.48%	1.65E-137
6	AP2	V_AP2ALPHA_02	3.47%	0
7	NRF	V_NRF2_Q4	3.01%	2.01E-140
8	NF1	V_NF1_Q6	2.96%	8.26E-146
9	EBOX	V_USF_Q6_01	2.84%	6.20E-121
10	VMAF	V_VMAF_01	2.41%	3.45E-101
11	MEF3	V_MEF3_B	2.38%	6.56E-95
12	PAX	V_PAX6_Q2	2.33%	2.13E-92
13	LMAF	V_LMAF_Q2	2.22%	4.71E-90
14	LRH1	V_LRH1_Q5	2.08%	1.51E-60
15	MYOGNF1	V_MYOGNF1_01	1.65%	3.20E-48
16	FXR	V_PXR_Q2	1.61%	3.00E-59
17	HIC1	V_HIC1_02	1.51%	1.41E-48
18	FOX	V_FREAC4_01	1.49%	5.06E-41
19	SP1	V_SP1_01	1.49%	2.58E-61
20	SMAD	V_SMAD4_Q6	1.41%	1.40E-46
21	MAF	V_MAF_Q6_01	1.40%	2.26E-64
22	AP4	V_AP4_Q6	1.33%	3.31E-40
23	DBP	V_DBP_Q6	1.30%	3.59E-28
24	ROAZ	V_ROAZ_01	1.16%	7.54E-39
25	MEIS1	V_MEIS1_01	0.94%	1.21E-26
26	E2F	V_E2F_Q4_01	0.88%	1.31E-28
27	HEN	V_LBP1_Q6	0.85%	2.79E-25
28	MIF1	V_MIF1_01	0.82%	3.15E-28
29	MYB	V_MYB_Q6	0.75%	1.17E-16
30	P53	V_P53_DECAMER_Q2	0.72%	3.30E-21
31	TGTGGT	V_PEBP_Q6	0.72%	1.59E-20
32	CACCC	V_CACCCBINDING_Q6	0.72%	9.66E-20
33	GATA	V_GATA1_05	0.71%	8.54E-12
34	CP2	V_CP2_02	0.68%	4.04E-38
35	SEF1	V_SEF1_C	0.63%	4.65E-14
36	SP3	V_SP3_Q3	0.56%	2.28E-23
37	ETS	V_CETS1P54_01	0.55%	5.51E-20
38	TGIF	V_TGIF_01	0.54%	1.77E-12
39	CMAF	V_CMAF_01	0.50%	4.34E-13
40	DEAF1	V_DEAF1_01	0.50%	8.08E-20
41	STAT	V_STAT1_01	0.47%	2.14E-18
42	DEC	V_DEC_Q1	0.46%	3.18E-12
43	OLF1	V_OLF1_01	0.44%	2.82E-17
44	TEF	V_TEF1_Q6	0.42%	0.041814963
45	ZF5	V_ZF5_B	0.41%	5.71E-18

**Table 4.5. (cont.)**

<b>Rank</b>	<b>Motif Family</b>	<b>Transfac</b>	<b>% of sites with motif</b>	<b>p-value</b>
46	RFX	V_RFX1_01	0.39%	4.19E-13
47	KAISO	V_KAISO_01	0.38%	2.87E-12
48	XPF1	V_XPF1_Q6	0.36%	4.70E-21
49	PTF1BETA	V_PTF1BETA_Q6	0.36%	5.06E-08
50	MINI	V_MINI19_B	0.31%	3.23E-18
51	EGR	V_EGR1_01	0.31%	1.30E-13
52	HES	V_HES1_Q2	0.30%	1.25E-12
53	E2	V_E2_01	0.29%	3.03E-17
54	SREB	V_SREBP_Q3	0.28%	6.50E-14
55	CAAT	V_YY1_02	0.26%	2.15E-14
56	GLI	V_ZIC3_01	0.26%	7.60E-15
57	SOX	V_SOX10_Q6	0.22%	1.46E-06
58	NANOG	V_NANOG_01	0.20%	4.48E-09
59	BRACH	V_TBX5_02	0.20%	6.29E-06
60	EBF	V_EBF_Q6	0.18%	1.08E-10
61	MTF1	V_MTF1_Q4	0.17%	2.89E-08
62	P300	V_P300_01	0.17%	2.98E-10
63	NFKB	V_NFMUE1_Q6	0.15%	2.20E-09
64	CACCT	V_AREB6_01	0.14%	1.87E-10
65	WT1	V_WT1_Q6	0.13%	2.11E-09
66	IK	V_IK1_01	0.06%	1.17E-08
67	R	V_R_01	0.06%	1.68E-05
68	ZBRK1	V_ZBRK1_01	0.06%	1.77E-06
69	GATA DIMER	V_EVI1_06	0.03%	0.002920699
70	AHR	V_AHR_01	0%	0.000133724

We used MotifEnrich to scan for enriched motifs in ER $\alpha$ [ $\alpha\beta$  cells] binding sites. The motifs are ranked according to their enrichment (descending order) in ER $\alpha$ [ $\alpha\beta$  cells] binding sites.

**Table 4.6.** TFBS enrichment in ER $\beta$ [ $\alpha\beta$  cells] binding sites.

Rank	Motif Family	Transfac ID	% of sites with motif	p-value
1	ERE	V_ER_Q6_Q2	16.80%	0
2	AP1	V_AP1_Q1	4.91%	0
3	BACH	V_BACH2_Q1	4.88%	0
4	CREB	V_ATF3_Q6	4.26%	0
5	EBOX	V_USF_Q6_Q1	3.84%	0
6	AP2	V_AP2ALPHA_Q1	3.64%	0
7	SP1	V_SP1_Q1	3.51%	0
8	E2F	V_E2F_Q6_Q1	3.41%	8.42E-128
9	NRF	V_NRF2_Q4	3.37%	0
10	HIC1	V_HIC1_Q2	3.30%	0
11	AR	V_AR_Q1	3.18%	3.08E-152
12	ZF5	V_ZF5_B	2.91%	4.78E-133
13	NF1	V_NF1_Q6	2.67%	1.59E-115
14	EGR	V_KROX_Q6	2.45%	1.34E-89
15	DEAF1	V_DEAF1_Q1	2.31%	6.97E-90
16	VMAF	V_VMAF_Q1	2.27%	3.54E-77
17	PAX	V_PAX5_Q1	2.25%	6.59E-103
18	LMAF	V_LMAF_Q2	2.21%	3.53E-93
19	MEF3	V_MEF3_B	2.21%	2.47E-81
20	ETS	V_GABP_B	2.12%	9.77E-72
21	LRH1	V_LRH1_Q5	2.04%	2.12E-62
22	SMAD	V_SMAD4_Q6	1.97%	1.89E-78
23	MOVO	V_MOVOB_Q1	1.77%	8.29E-67
24	MYOGNF1	V_MYOGNF1_Q1	1.70%	1.19E-53
25	ROAZ	V_ROAZ_Q1	1.67%	2.84E-64
26	MINI	V_MINI19_B	1.61%	5.16E-58
27	HEN	V_HEN1_Q2	1.61%	6.67E-62
28	MAF	V_MAF_Q6_Q1	1.53%	7.60E-61
29	FXR	V_PXR_Q2	1.46%	1.25E-51
30	AP4	V_AP4_Q6	1.44%	1.15E-52
31	MIF1	V_MIF1_Q1	1.34%	4.29E-45
32	CP2	V_CP2_Q2	1.34%	1.16E-49
33	CACCC	V_CACCCBINDING_Q6	1.18%	3.81E-37
34	AHR	V_AHRARNT_Q2	1.16%	1.23E-41
35	WT1	V_WT1_Q6	1.05%	2.00E-37
36	GGG	V_CHCH_Q1	1%	1.54E-35
37	OLF1	V_OLF1_Q1	0.95%	1.32E-30
38	MEIS1	V_MEIS1_Q1	0.87%	4.32E-26
39	WHN	V_WHN_B	0.84%	7.76E-32
40	STAT	V_STAT1_Q1	0.83%	3.87E-32
41	MYB	V_CMYB_Q1	0.83%	2.54E-30
42	SP3	V_SP3_Q3	0.79%	1.23E-32
43	HES	V_HES1_Q2	0.79%	5.66E-28
44	NFKB	V_NFKAPPAB50_Q1	0.76%	4.21E-27
45	P53	V_P53_Q1	0.72%	7.43E-16



**Table 4.6. (cont.)**

<b>Rank</b>	<b>Motif Family</b>	<b>Transfac ID</b>	<b>% of sites with motif</b>	<b>p-value</b>
46	E2	V_E2_Q6	0.72%	5.03E-28
47	RFX	V_RFX1_01	0.64%	2.78E-17
48	DBP	V_DBP_Q6	0.62%	3.38E-11
49	MTF1	V_MTF1_Q4	0.55%	7.08E-23
50	FOX	V_FREAC4_01	0.54%	1.47E-08
51	TGTGGT	V_PEBP_Q6	0.53%	4.93E-15
52	GLI	V_ZIC3_01	0.51%	4.46E-16
53	R	V_R_01	0.50%	1.41E-17
54	DEC	V_DEC_Q1	0.48%	1.30E-15
55	ZNF219	V_ZNF219_01	0.43%	7.52E-22
56	CMAF	V_CMAF_01	0.43%	1.96E-12
57	P300	V_P300_01	0.38%	8.54E-16
58	VMYB	V_VMYB_02	0.38%	8.79E-15
59	KAISO	V_KAISO_01	0.36%	1.11E-16
60	XPF1	V_XPF1_Q6	0.33%	2.70E-13
61	CAAT	V_NFY_01	0.30%	7.37E-08
62	SREB	V_SREBP_Q3	0.28%	4.87E-12
63	PTF1BETA	V_PTF1BETA_Q6	0.26%	1.12E-09
64	NANOG	V_NANOG_01	0.26%	6.91E-11
65	TGIF	V_TGIF_01	0.23%	5.46E-10
66	ZBRK1	V_ZBRK1_01	0.23%	1.86E-10
67	ZID	V_ZID_01	0.22%	1.39E-12
68	EBF	V_EBF_Q6	0.20%	2.72E-12
69	GCM	V_GCM_Q2	0.19%	1.23E-10
70	LRF	V_LRF_Q2	0.18%	5.55E-14
71	CACCT	V_AREB6_01	0.12%	1.47E-09
72	SPZ	V_SPZ1_01	0.09%	5.83E-13
73	SEF1	V_SEF1_C	0.08%	3.34E-06
74	TFIII	V_TFIII_Q6	0.07%	3.32E-11
75	SOX	V_SOX10_Q6	0.04%	0.016002141
76	BRACH	V_TBX5_02	0.03%	0.0002267
77	HMX1	V_HMX1_01	0.01%	2.27E-08
78	STAF	V_STAF_01	0%	1.06E-05

We used MotifEnrich to scan for enriched motifs in ER $\beta$ [ $\alpha\beta$  cells] binding sites. The motifs are ranked according to their enrichment (descending order) in ER $\beta$ [ $\alpha\beta$  cells] binding sites.

**Table 4.7.** TFBS enrichment in ER $\beta$ [ $\beta$  cells] binding sites.

Rank	Motif Family	Transfac ID	% of sites with motif	p-value
1	ERE	V_ER_Q6_Q2	16.36%	0
2	AP1	V_AP1_Q6	4.53%	0
3	BACH	V_BACH2_Q1	4.36%	0
4	CREB	V_ATF3_Q6	4.19%	0
5	AP2	V_AP2ALPHA_Q2	3.91%	0
6	EBOX	V_USF_Q6_Q1	3.66%	0
7	AR	V_AR_Q1	3.37%	0
8	NRF	V_NRF2_Q4	3.05%	0
9	LMAF	V_LMAF_Q2	2.58%	0
10	HIC1	V_HIC1_Q2	2.50%	0
11	LRH1	V_LRH1_Q5	2.42%	2.80E-156
12	NF1	V_NF1_Q6	2.36%	0
13	MEF3	V_MEF3_B	2.36%	0
14	PAX	V_PAX6_Q2	2.27%	5.18E-114
15	VMAF	V_VMAF_Q1	2.11%	1.75E-114
16	SP1	V_SP1_Q1	2.08%	5.19E-130
17	SMAD	V_SMAD4_Q6	2.07%	1.87E-124
18	E2F	V_E2F_Q4_Q1	1.84%	1.57E-91
19	FXR	V_PXR_Q2	1.83%	1.08E-128
20	MYOGNF1	V_MYOGNF1_Q1	1.64%	1.15E-82
21	ROAZ	V_ROAZ_Q1	1.56%	5.78E-91
22	DEAF1	V_DEAF1_Q1	1.53%	6.24E-80
23	MAF	V_MAF_Q6_Q1	1.49%	9.43E-96
24	HEN	V_HEN1_Q2	1.49%	2.47E-85
25	ZF5	V_ZF5_B	1.44%	4.03E-66
26	AP4	V_AP4_Q5	1.36%	8.55E-89
27	CP2	V_CP2_Q2	1.35%	1.50E-75
28	MIF1	V_MIF1_Q1	1.34%	6.54E-66
29	MEIS1	V_MEIS1_Q1	1.29%	3.04E-60
30	CACCC	V_CACCCFACTOR_Q6	1.25%	2.63E-62
31	ETS	V_NERF_Q2	1.21%	7.12E-73
32	EGR	V_EGR3_Q1	1.14%	2.56E-51
33	MYB	V_CMYB_Q1	1.12%	2.03E-49
34	MINI	V_MINI19_B	1.12%	4.75E-59
35	P53	V_P53_Q2	1.06%	3.99E-54
36	DBP	V_DBP_Q6	1.01%	1.16E-30
37	FOX	V_FREAC4_Q1	1%	2.02E-27
38	SP3	V_SP3_Q3	0.91%	1.67E-47
39	OLF1	V_OLF1_Q1	0.80%	2.81E-40
40	HES	V_HES1_Q2	0.77%	5.87E-41
41	EBF	V_EBF_Q6	0.76%	7.60E-36
42	STAT	V_STAT1_Q1	0.73%	1.36E-39
43	TGIF	V_TGIF_Q1	0.71%	6.60E-19
44	TGTGGT	V_PEBP_Q6	0.69%	1.24E-25

**Table 4.7. (cont.)**

<b>Rank</b>	<b>Motif Family</b>	<b>Transfac ID</b>	<b>% of sites with motif</b>	<b>p-value</b>
45	WT1	V_WT1_Q6	0.58%	1.85E-28
46	MOVO	V_MOVOB_01	0.58%	1.84E-34
47	AHR	V_AHRARNT_02	0.56%	3.71E-26
48	MTF1	V_MTF1_Q4	0.53%	1.96E-25
49	E2	V_E2_01	0.53%	3.26E-37
50	CMAF	V_CMAF_01	0.53%	7.60E-14
51	XPF1	V_XPF1_Q6	0.52%	6.37E-19
52	NFKB	V_NFKAPPAB50_01	0.52%	3.44E-24
53	WHN	V_WHN_B	0.52%	4.15E-29
54	RFX	V_RFX1_01	0.51%	6.71E-22
55	GGG	V_CHCH_01	0.48%	5.86E-26
56	P300	V_P300_01	0.45%	2.18E-28
57	PTF1BETA	V_PTF1BETA_Q6	0.43%	1.13E-14
58	DEC	V_DEC_Q1	0.43%	1.95E-21
59	CAAT	V_YY1_02	0.40%	3.78E-20
60	CACCT	V_AREB6_01	0.37%	6.42E-18
61	BRACH	V_TBX5_02	0.37%	9.35E-14
62	GLI	V_ZIC3_01	0.36%	9.58E-31
63	SREB	V_SREBP_Q3	0.35%	1.70E-18
64	KAISO	V_KAISO_01	0.35%	1.63E-22
65	SEF1	V_SEF1_C	0.32%	1.33E-12
66	R	V_R_01	0.31%	7.79E-19
67	ZBRK1	V_ZBRK1_01	0.29%	6.01E-13
68	ZNF219	V_ZNF219_01	0.22%	2.83E-20
69	ZID	V_ZID_01	0.22%	6.31E-16
70	NANOG	V_NANOG_01	0.18%	8.59E-15
71	IK	V_IK1_01	0.18%	1.56E-08
72	TFIII	V_TFIII_Q6	0.15%	1.68E-09
73	VMYB	V_VMYB_02	0.15%	3.45E-11
74	TEF	V_TEF1_Q6	0.13%	0.871926
75	GATA	V_GATA1_05	0.08%	0.23605
76	GCM	V_GCM_Q2	0.08%	2.32E-12
77	SZF11	V_SZF11_01	0.06%	0.001904
78	HMX1	V_HMX1_01	0.04%	7.01E-10
79	SOX	V_SOX10_Q6	0.01%	0.147899
80	IPF	V_IPF1_Q4_01	0.01%	0.00013

We used MotifEnrich to scan for enriched motifs in ER $\beta$ [ $\beta$  cells] binding sites. The motifs are ranked according to their enrichment (descending order) in ER $\beta$ [ $\beta$  cells] binding sites.

**Table 4.8.** Evaluation of E2F motifs and their enrichment at ER binding sites.

**A.** Enrichment of E2F motif at ER $\alpha$  [ $\alpha$  cells] and ER $\beta$  [ $\beta$  cells] binding sites

<b>ER binding sites</b>	<b>% of ER sites with E2F motif</b>	<b>p-value</b>
ER $\alpha$ [ $\alpha$ cells]	0.51%	1.54E-33
ER $\beta$ [ $\beta$ cells]	1.74%	6.90E-87

**B.** Enrichment of E2F motif at ER $\alpha$  and ER $\beta$  binding sites in [ $\alpha\beta$  cells]

<b>ER binding sites</b>	<b>% of ER sites with E2F motif</b>	<b>p-value</b>
ER $\alpha$ -only [ $\alpha\beta$ cells]	0%	1
ER $\beta$ -only [ $\alpha\beta$ cells]	6.89%	8.25E-160
ER $\alpha$ [ $\alpha\beta$ cells] & ER $\beta$ [ $\alpha\beta$ cells]	0.85%	3.29E-25

**Table 4.9.** Summary of SRC3 and RIP140 chromatin affinity sites in the three MCF-7 cells expressing ER $\alpha$  only, both ER $\alpha$  and ER $\beta$ , or ER $\beta$  only.

	<b>[<math>\alpha</math> cells]</b>	<b>[<math>\alpha\beta</math> cells]</b>	<b>[<math>\beta</math> cells]</b>
Total SRC3 binding sites	4030	6833	6056
Total RIP140 binding sites	2119	2513	8924

**Table 4.10.** Segregation of ER binding sites arrays.

	[ $\alpha$ cells]	[ $\alpha\beta$ cells]	[ $\alpha\beta$ cells]	[ $\beta$ cells]
<b>E+S+R</b>	1737 sites ER $\alpha$ [ $\alpha$ ]+SRC3[ $\alpha$ ]+ RIP140[ $\alpha$ ]	2266 sites ER $\alpha$ [ $\alpha\beta$ ]+SRC3[ $\alpha\beta$ ]+ RIP140[ $\alpha\beta$ ]	2257 sites ER $\beta$ [ $\alpha\beta$ ]+SRC3[ $\alpha\beta$ ]+ RIP140[ $\alpha\beta$ ]	5047 sites ER $\beta$ [ $\beta$ ]+SRC3[ $\beta$ ]+ RIP140[ $\beta$ ]
<b>E+S</b>	1834 sites ER $\alpha$ [ $\alpha$ ]+SRC3[ $\alpha$ ]	3972 sites ER $\alpha$ [ $\alpha\beta$ ]+SRC3[ $\alpha\beta$ ]	3896 sites ER $\beta$ [ $\alpha\beta$ ]+SRC3[ $\alpha\beta$ ]	710 sites ER $\beta$ [ $\beta$ ]+SRC3[ $\beta$ ]
<b>E+R</b>	70 sites ER $\alpha$ [ $\alpha$ ]+RIP140[ $\alpha$ ]	41 sites ER $\alpha$ [ $\alpha\beta$ ]+RIP140[ $\alpha\beta$ ]	48 sites ER $\beta$ [ $\alpha\beta$ ]+RIP140[ $\alpha\beta$ ]	3294 sites ER $\beta$ [ $\beta$ ]+RIP140[ $\beta$ ]
<b>E</b>	34257 sites ER $\alpha$ [ $\alpha$ ]	13671 sites ER $\alpha$ [ $\alpha\beta$ ]	14027 sites ER $\beta$ [ $\alpha\beta$ ]	28706 sites ER $\beta$ [ $\beta$ ]

We separated ER binding sites from their respectively cell-types into 4 category, E+S+R, E+S, E+R, and E. ER binding sites that are co-localized with SRC3 and RIP140 are placed in E+S+R category. Category E+S are ER binding sites that are co-localized with SRC3 whereas Category E+R are ER binding sites that are co-localized with RIP140. The remaining ER binding sites that do not overlap SRC3 or RIP140 binding are placed in Category E. SRC3 and RIP140 ChIP-seq peaks that are within  $\pm 200$ bp from ER ChIP-seq peaks are defined as co-localized to the ER.

**Table 4.11.** Defining ER target genes.

	I* [ $\alpha$ cells]		II* [ $\alpha\beta$ cells]		III* [ $\alpha\beta$ cells]		IV* [ $\beta$ cells]	
	E2-induced genes	E2-repressed genes	E2-induced genes	E2-repressed genes	E2-induced genes	E2-repressed genes	E2-induced genes	E2-repressed genes
<b>E+S+R</b>	101	10	104	6	105	6	275	26
<b>E+S</b>	52	26	119	28	120	29	21	9
<b>E+R</b>	9	1	1	0	1	0	140	31
<b>E</b>	607	690	299	113	459	239	766	307

\*I: Overlapping of E2 regulated genes in [ $\alpha$  cells] with ER $\alpha$ [ $\alpha$  cells] binding sites from categories E+S+R, E+S, E+R, and E

\*II: Overlapping of E2 regulated genes in [ $\alpha\beta$  cells] with ER $\alpha$ [ $\alpha\beta$  cells] binding sites from categories E+S+R, E+S, E+R, and E

\*III: Overlapping of E2 regulated genes in [ $\alpha\beta$  cells] with ER $\beta$ [ $\alpha\beta$  cells] binding sites from categories E+S+R, E+S, E+R, and E

\*IV: Overlapping of E2 regulated genes in [ $\beta$  cells] with ER $\beta$ [ $\beta$  cells] binding sites from categories E+S+R, E+S, E+R, and E

E2-regulated genes were overlapped with ER binding sites from E+S+R, E+S, E+R, and E for each cell-type. Genes that have a ER binding site bound within their regulatory regions ( $\pm$  20kbp from the TSS) are identified and grouped according to the ER binding sites category.

**Table 4.12.** Analysis for enriched biological processes in ER $\alpha$  and ER $\beta$  target genes.

**A.** 10 most enriched biological processes terms of ER $\alpha$  [ $\alpha$  cells] targeted genes

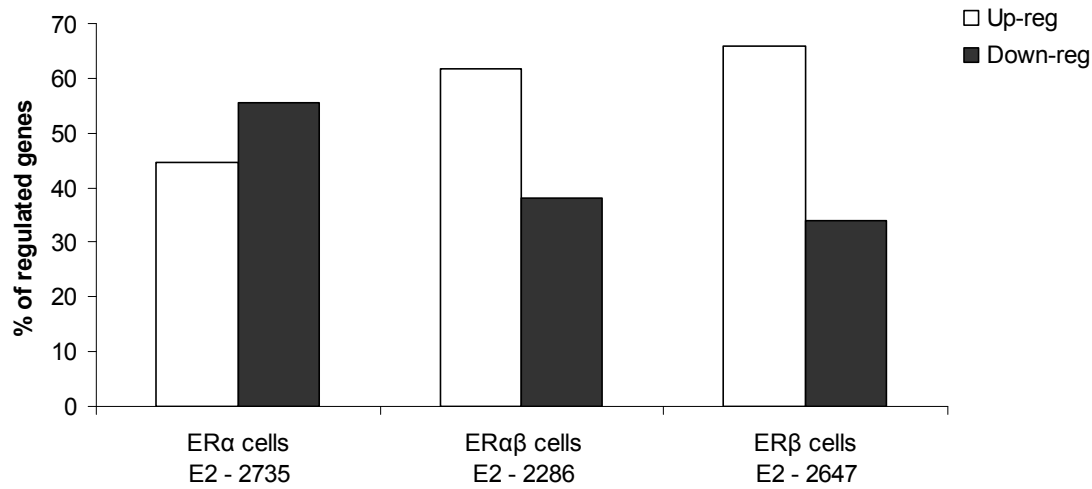
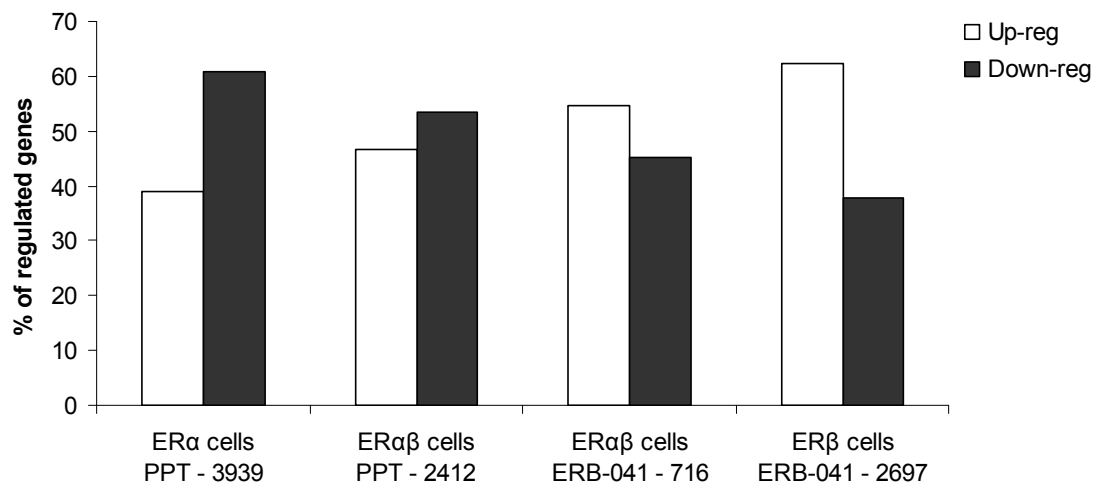
Rank	Term	Count	PValue	Fold Enrichment
1	GO:0042127~regulation of cell proliferation	16	0.001799943	2.460670194
2	GO:0008284~positive regulation of cell proliferation	10	0.006605061	2.937049942
3	GO:0006357~regulation of transcription from RNA polymerase II promoter	13	0.014552865	2.183330013
4	GO:0050678~regulation of epithelial cell proliferation	4	0.019138314	6.980814355
5	GO:0010629~negative regulation of gene expression	10	0.021120574	2.422918463
6	GO:0016055~Wnt receptor signaling pathway	5	0.022851186	4.596146856
7	GO:0032553~ribonucleotide binding	25	0.023905015	1.554364865
8	GO:0032555~purine ribonucleotide binding	25	0.023905015	1.554364865
9	GO:0030695~GTPase regulator activity	9	0.024941106	2.528695854
10	GO:0045892~negative regulation of transcription, DNA-dependent	8	0.025639154	2.744593678

**B.** 10 most enriched biological processes terms of ER $\beta$  [ $\beta$  cells] targeted genes

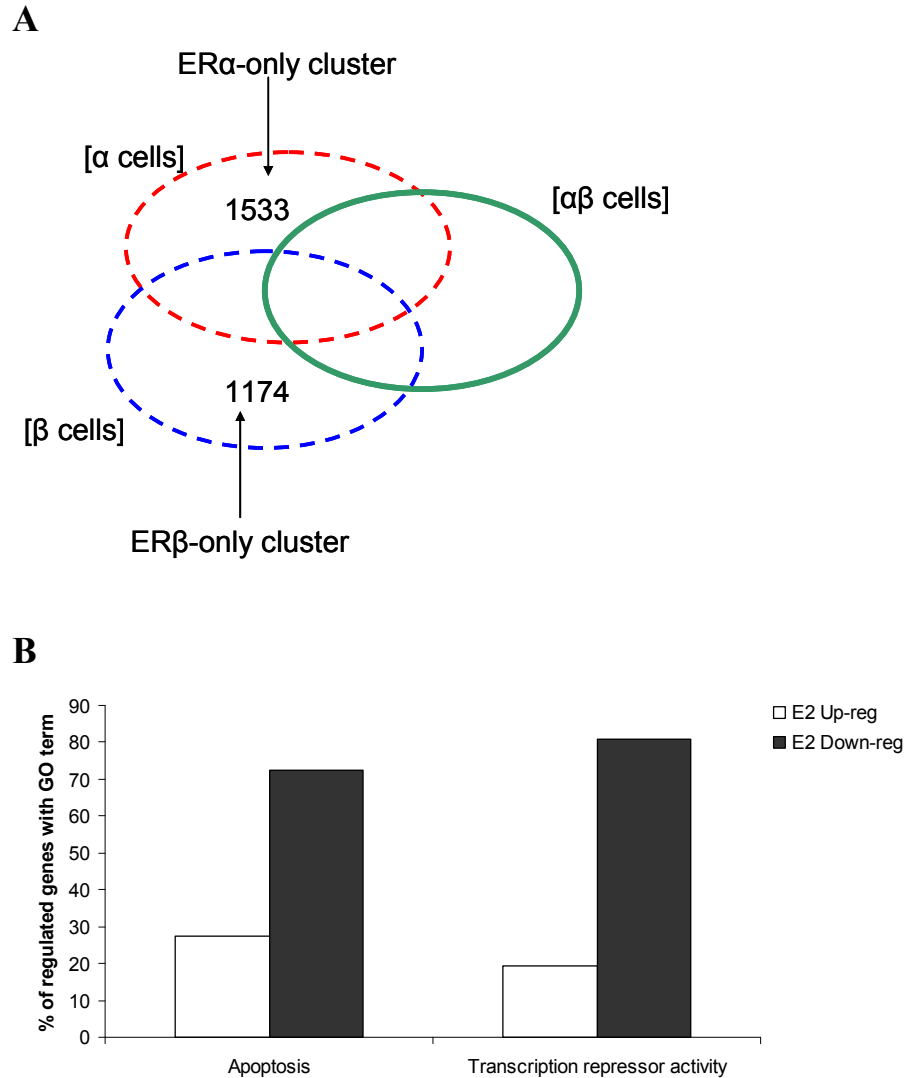
Rank	Term	Count	PValue	Fold Enrichment
1	GO:0042981~regulation of apoptosis	40	9.98E-06	2.130358963
2	GO:0043067~regulation of programmed cell death	40	1.25E-05	2.1091085
3	GO:0010941~regulation of cell death	40	1.38E-05	2.101248468
4	GO:0033273~response to vitamin	9	1.52E-04	5.766494375
5	GO:0043065~positive regulation of apoptosis	24	1.84E-04	2.388007082
6	GO:0012501~programmed cell death	30	1.90E-04	2.117911123
7	GO:0043068~positive regulation of programmed cell death	24	2.05E-04	2.371268715
8	GO:0010942~positive regulation of cell death	24	2.19E-04	2.360239558
9	GO:0006915~apoptosis	29	3.36E-04	2.0785443
10	GO:0033189~response to vitamin A	7	4.18E-04	7.04793757

DAVID analysis for enriched biological processes terms of genes whose regulatory regions are bound by ER $\alpha$  with SRC3 and RIP140 in [ $\alpha$  cells] and genes that are regulatory regions are bound by ER $\beta$  with SRC3 and RIP140 in [ $\beta$  cells]. Top 10 most enriched biological processes terms as identified by DAVID are shown.



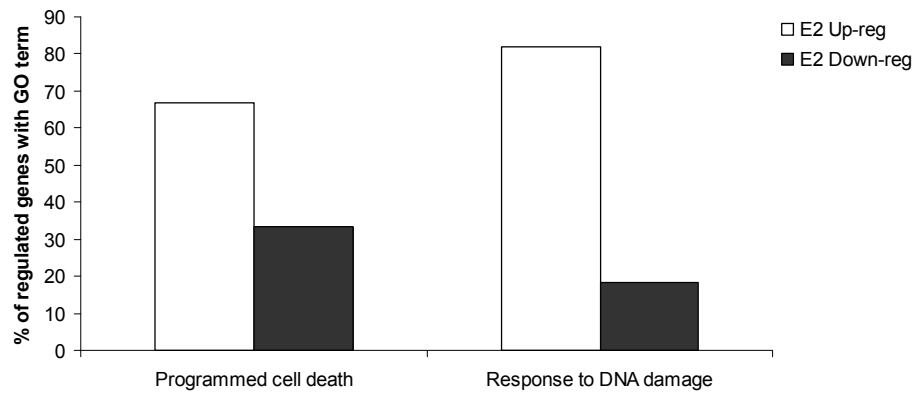
**A****B**

**Figure 4.1.** Gene Regulation in MCF-7 Cells Containing Various Complements of ER $\alpha$  and ER $\beta$ . (A) Percentage of up- and down- regulated genes upon E2 treatment in cells expressing ER $\alpha$  alone, ER $\alpha$  and ER $\beta$ , and ER $\beta$  alone. (B) Percentage of up- and down- regulated genes upon ER sub-type selective ligands treatment in cells expressing ER $\alpha$  alone, ER $\alpha$  and ER $\beta$ , and ER $\beta$  alone. PPT is an ER $\alpha$  selective ligand whereas ERB-041 only activates ER $\beta$ .

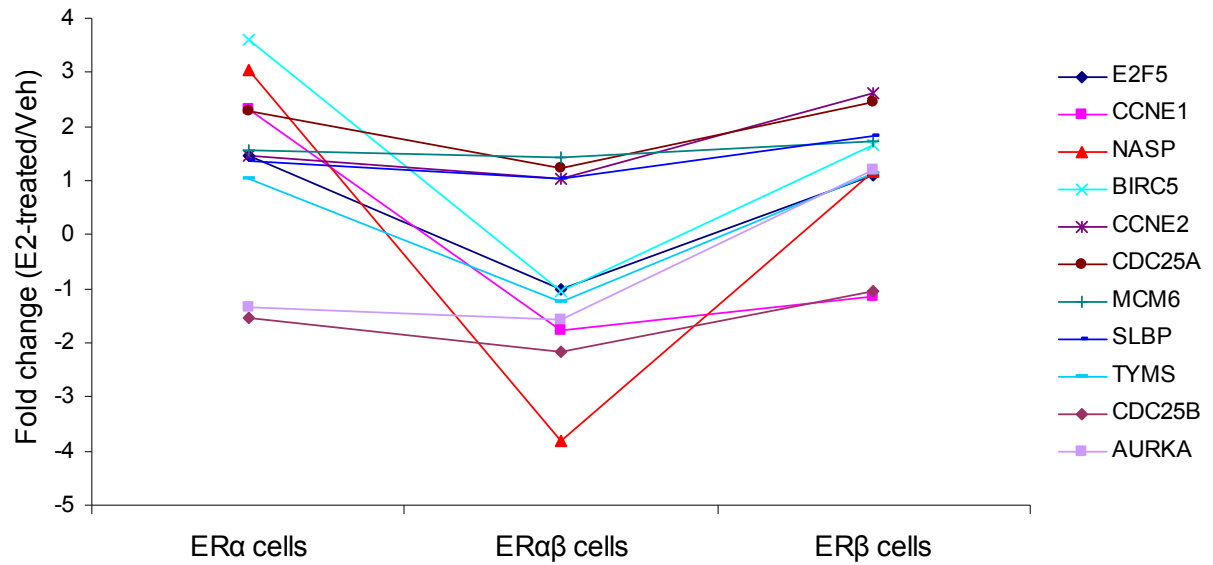
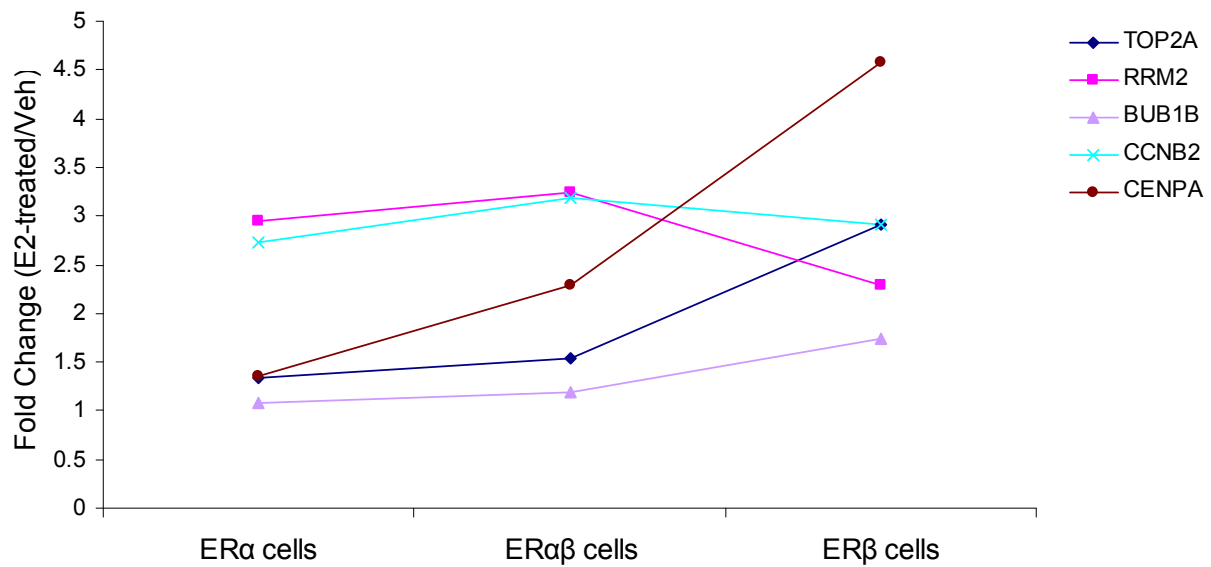


**Figure 4.2.** Global Analysis of Cellular Processes Regulated Uniquely by ER $\alpha$  and ER $\beta$ . (A) Venn diagram showing the subsets of E2-regulated genes used for GO analysis. Only genes that are uniquely regulated in each cell-type were selected. (B) In the ER $\alpha$ -only cluster, we observed that most of the genes with enriched GO term of “Apoptosis” and “Transcription repressor activity” are E2-repressed. (C) Genes with enriched GO term of “Programmed cell death” and “Response to DNA damage” are mostly E2-induced in the ER $\beta$ -only cluster.

**C**

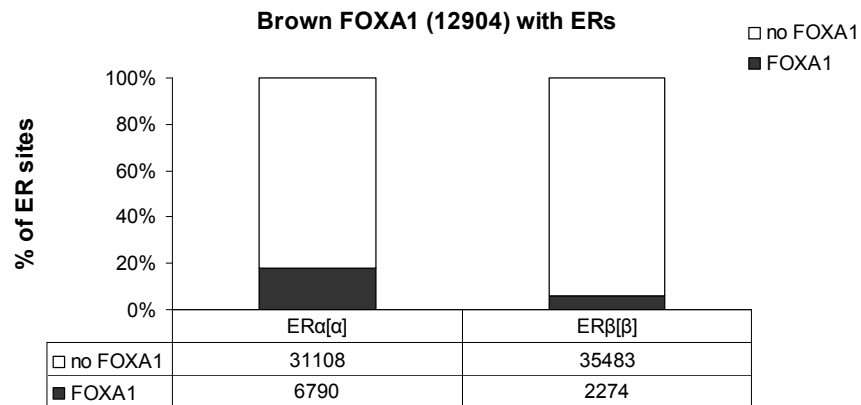


**Figure 4.2.** (cont.)

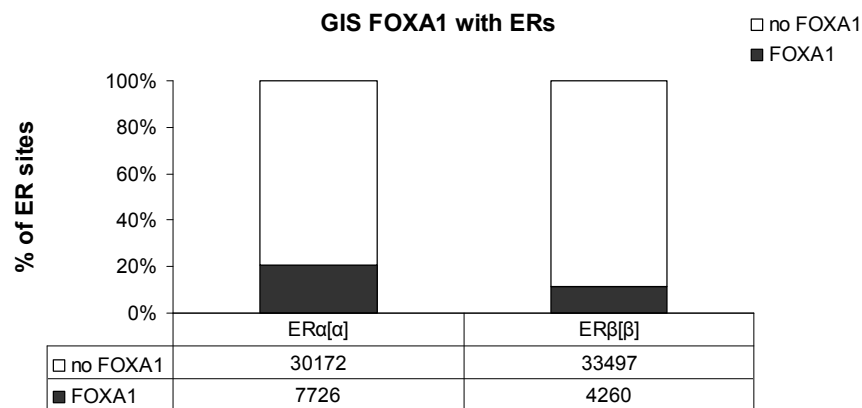
**A****B**

**Figure 4.3.** Gene expression profiles of cell cycle and proliferation genes. The 19 cell cycle and proliferation genes were examined for their E2-incuded expression in the microarray data. (A) Genes are down-modulated by ER $\beta$  (comparing expression to ER $\alpha$  cells) in the presence of ER $\alpha$  (B) Genes up-regulated by ER $\alpha$  and ER $\beta$ .

**A**

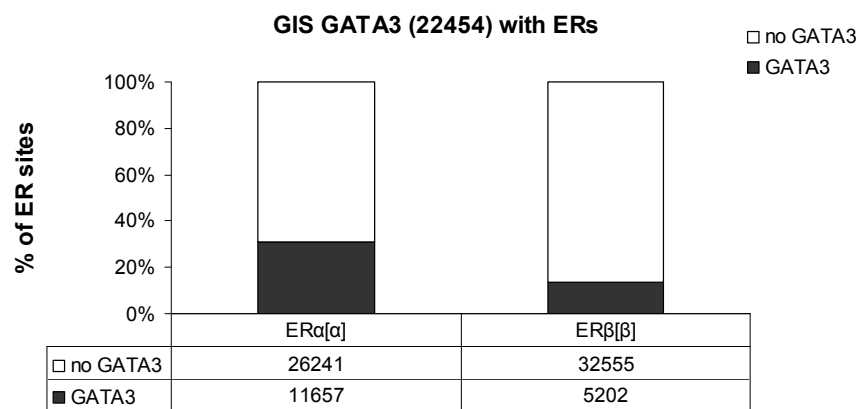


**B**



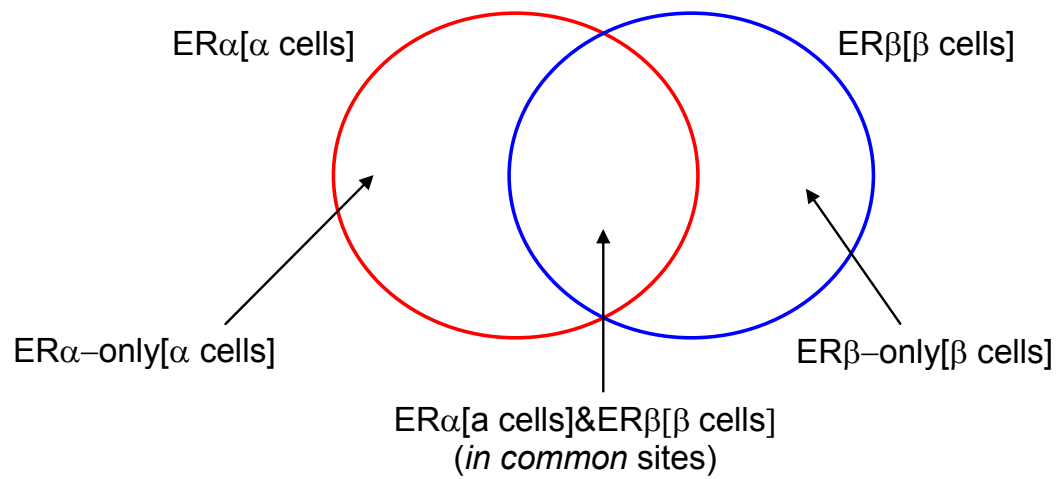
**Figure 4.4.** Genome-wide analysis of FOXA1 and GATA3 binding sites with ERα[α cells] and ERβ[β cells] binding sites. We analyzed the overlap of ERα[α cells] and ERβ[β cells] binding sites for FOXA1 binding sites and GATA3 binding sites. (A) We overlapped ERα[α cells] and ERβ[β cells] binding sites with FOXA1 ChIP-chip binding sites [47]. (B) ERα[α cells] and ERβ[β cells] binding sites were overlapped with FOXA1 ChIP-seq binding sites (GIS unpublished data, Edison Liu). (C) ERα[α cells] and ERβ[β cells] binding sites were overlapped with GATA3 ChIP-seq binding sites (GIS unpublished data, Edison Liu).

C

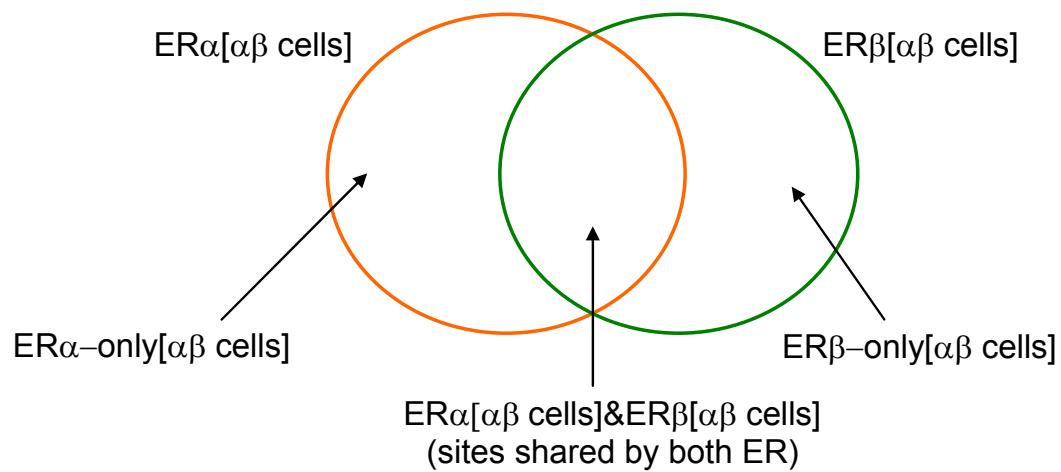


**Figure 4.4.** (cont.)

**A**

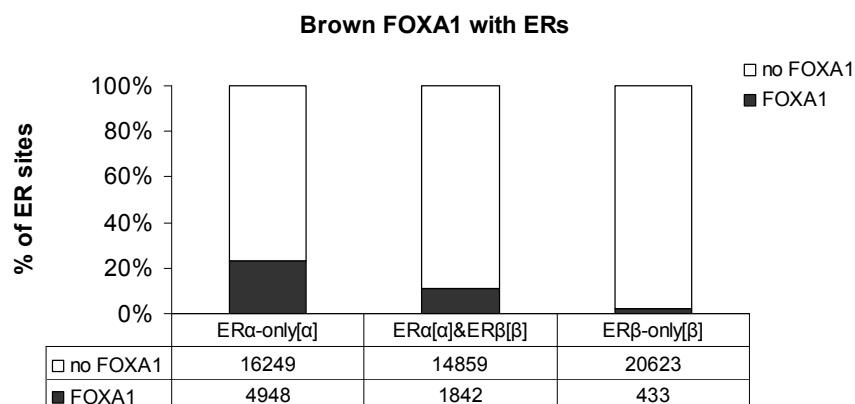


**B**

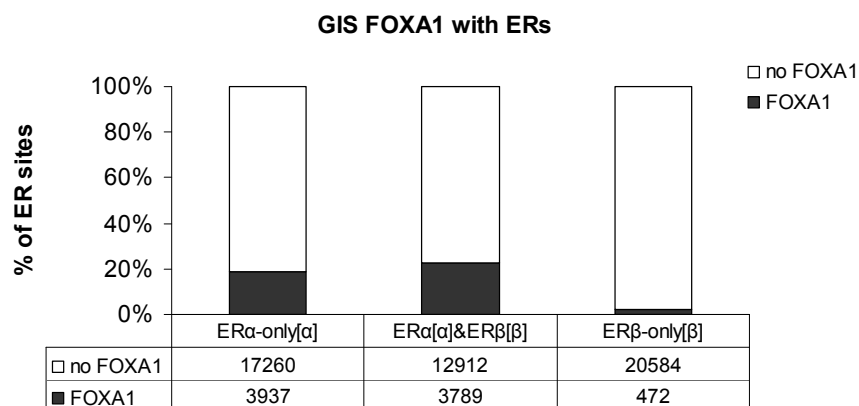


**Figure 4.5.** Venn diagrams showing the overlap of ER binding sites when they are present either separately or together in MCF-7 cells. (A) Overlap of  $ER\alpha$  [ $\alpha$  cells] with  $ER\beta$  [ $\beta$  cells] binding sites. (B) Overlap of  $ER\alpha$  [ $\alpha\beta$  cells] with  $ER\beta$  [ $\alpha\beta$  cells] binding sites.

**A**



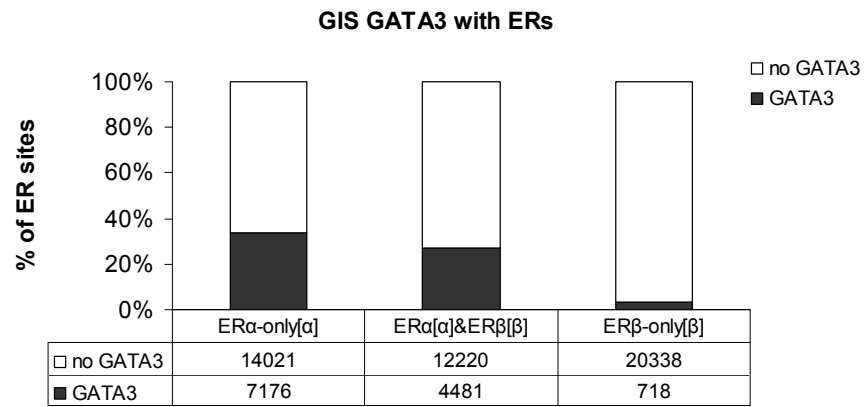
**B**



**Figure 4.6.** Genome-wide analysis of FOXA1 and GATA3 binding sites with ERα[α cells] and ERβ[β cells] binding sites. ERα [α cells] and ERβ [β cells] binding sites are separated into ERα-only [α cells], ERβ-only [β cells], and sites *in common* to both ERs (ERα[α cells]&ERβ[β cells]). These ER sites are then reanalyzed for overlap with FOXA1 and GATA3 binding. (A) Overlap of ERs binding sites with FOXA1 ChIP-chip binding sites [47]. (B) ERs binding sites were overlapped with FOXA1 ChIP-seq binding sites (GIS unpublished data, Edison Liu). (C) ERs binding sites were overlapped with GATA3 ChIP-seq binding sites (GIS unpublished data, Edison Liu).

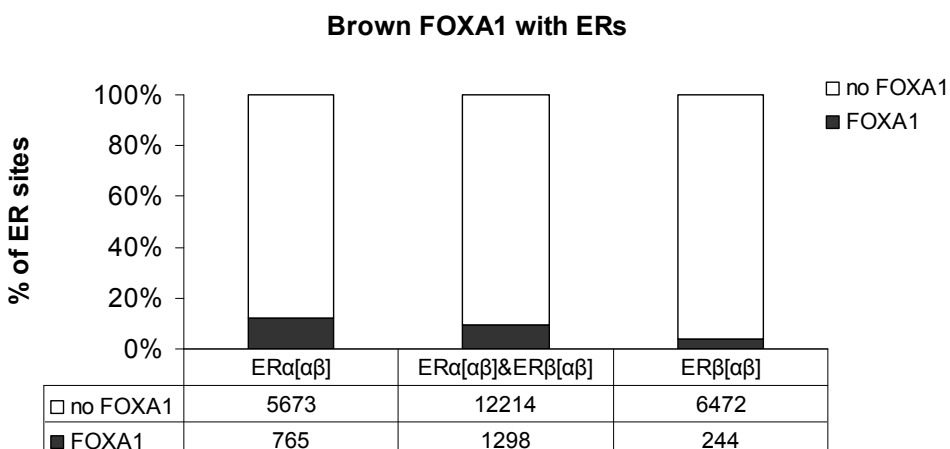


C

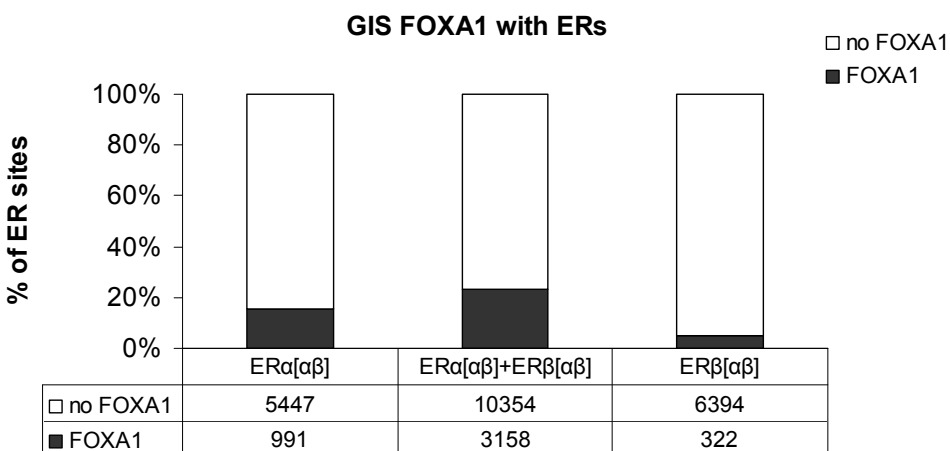


**Figure 4.6.** (cont.)

A

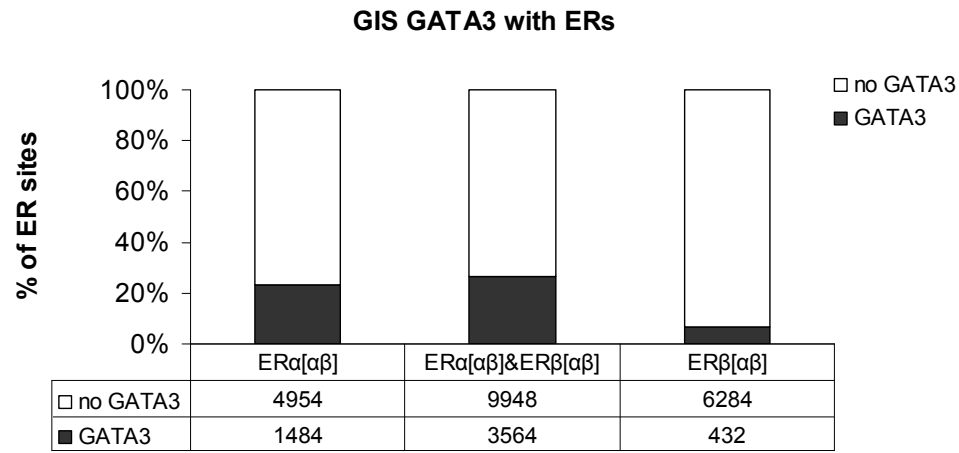


B

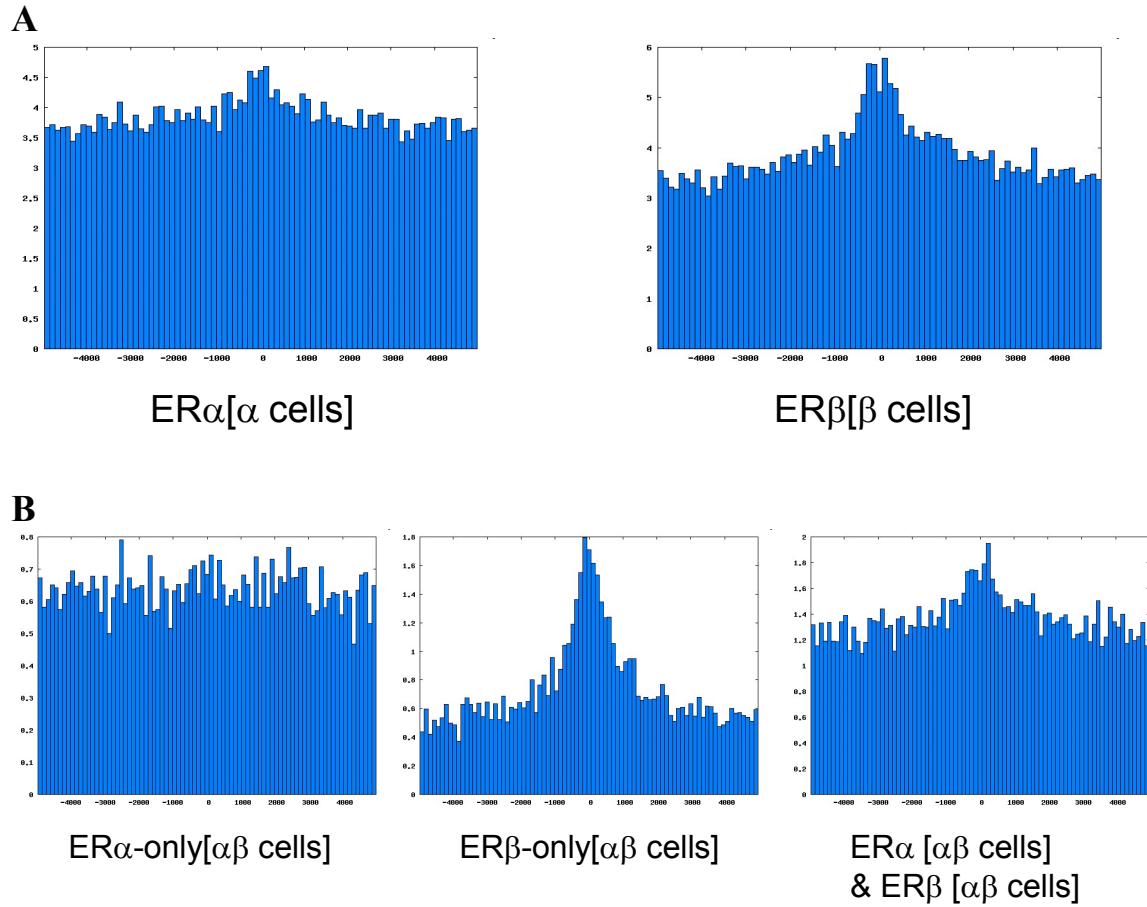


**Figure 4.7.** Genome-wide analysis of FOXA1 and GATA3 binding sites with ERα[αβ cells] and ERβ[αβ cells] binding sites. ERα [αβ cells] and ERβ [αβ cells] binding sites are separated to sites where ERα can uniquely binds (ERα-only[αβ cells]), ERβ can uniquely binds (ERβ-only[αβ cells], and where both ER can binds (ERα [αβ cells]&ERβ[αβ cells])). FOXA1 and GATA3 binding site are then overlapped with these sites. (A) Overlap of ERs binding sites with FOXA1 ChIP-chip binding sites [47]. (B) ERs binding sites were overlapped with FOXA1 ChIP-seq binding sites (GIS unpublished data, Edison Liu). (C) ERs binding sites were overlapped with GATA3 ChIP-seq binding sites (GIS unpublished data, Edison Liu).

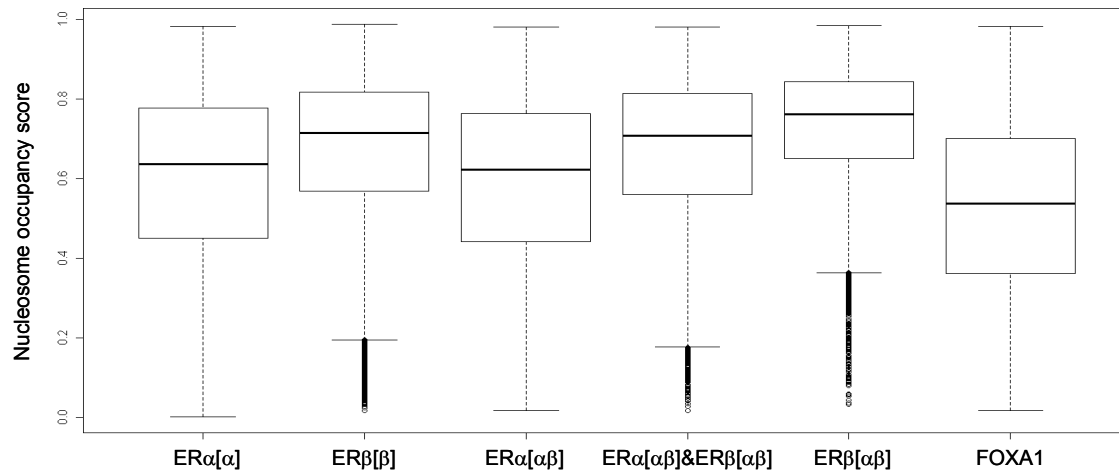
C



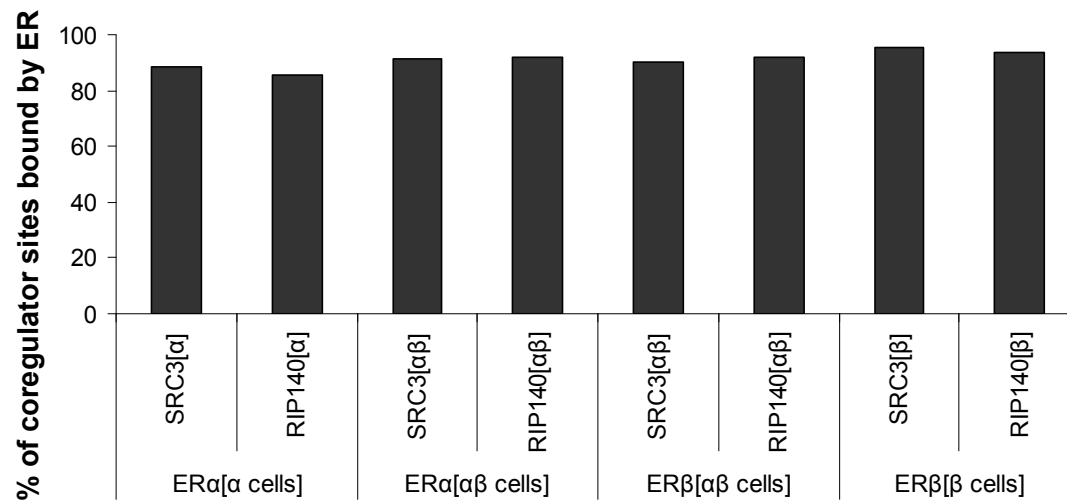
**Figure 4.7.** (cont.)



**Figure 4.8.** Histograms of the occurrences of E2F motif. The frequency of occurrences of the E2F motif centered on ER binding sites is plotted. We extend the region to 5000bp up- and down-stream of the ER peak. (A) E2F motif occurrences around ER $\alpha$  [ $\alpha$  cells] binding sites and ER $\beta$  [ $\beta$  cells] binding sites. (B) E2F motif occurrences around ER $\alpha$ -only binding sites, ER $\beta$ -only binding sites, and ER $\alpha$  & ER $\beta$  binding sites in [ $\alpha\beta$  cells].



**Figure 4.9.** Intrinsic Nucleosome Occupancy across ER sites. We used the intrinsic nucleosome preference model (proposed by Kaplan et al. [51]) to estimate nucleosome occupancy at ER $\alpha$  and ER $\beta$  sites. As reference, we also considered the nucleosome occupancy of FOXA1 sites. A score of 1 will indicate full intrinsic nucleosome occupancy whereas a score of 0 will indicate nucleosome free region.



**Figure 4.10.** Co-localization of ER binding sites with SRC3 and RIP140. We overlapped SRC3 and RIP140 ChIP-seq data to ER binding sites in each of the respective cell-type (in [α cells], [αβ cells], and [β cells]). The majority of SRC3 and RIP140 are co-localized with ER binding sites and they are recruited equally well by both ERα and ERβ.

#### 4.7 REFERENCES

1. Deroo, B.J. and K.S. Korach, *Estrogen receptors and human disease*. J Clin Invest, 2006. **116**(3): p. 561-70.
2. Nilsson, S., et al., *Mechanisms of estrogen action*. Physiol Rev, 2001. **81**(4): p. 1535-65.
3. Frasor, J., et al., *Profiling of estrogen up- and down-regulated gene expression in human breast cancer cells: insights into gene networks and pathways underlying estrogenic control of proliferation and cell phenotype*. Endocrinology, 2003. **144**(10): p. 4562-74.
4. Chang, E.C., et al., *Estrogen Receptors alpha and beta as determinants of gene expression: influence of ligand, dose, and chromatin binding*. Mol Endocrinol, 2008. **22**(5): p. 1032-43.
5. Lazennec, G., et al., *ER beta inhibits proliferation and invasion of breast cancer cells*. Endocrinology, 2001. **142**(9): p. 4120-30.
6. Paruthiyil, S., et al., *Estrogen receptor beta inhibits human breast cancer cell proliferation and tumor formation by causing a G2 cell cycle arrest*. Cancer Res, 2004. **64**(1): p. 423-8.
7. Strom, A., et al., *Estrogen receptor beta inhibits 17beta-estradiol-stimulated proliferation of the breast cancer cell line T47D*. Proc Natl Acad Sci U S A, 2004. **101**(6): p. 1566-71.
8. Williams, C., et al., *A genome-wide study of the repressive effects of estrogen receptor beta on estrogen receptor alpha signaling in breast cancer cells*. Oncogene, 2008. **27**(7): p. 1019-32.
9. Lin, C.Y., et al., *Inhibitory effects of estrogen receptor beta on specific hormone-responsive gene expression and association with disease outcome in primary breast cancer*. Breast Cancer Res, 2007. **9**(2): p. R25.
10. Chang, E.C., et al., *Impact of estrogen receptor beta on gene networks regulated by estrogen receptor alpha in breast cancer cells*. Endocrinology, 2006. **147**(10): p. 4831-42.
11. Glass, C.K. and M.G. Rosenfeld, *The coregulator exchange in transcriptional functions of nuclear receptors*. Genes Dev, 2000. **14**(2): p. 121-41.
12. McKenna, N.J. and B.W. O'Malley, *Combinatorial control of gene expression by nuclear receptors and coregulators*. Cell, 2002. **108**(4): p. 465-74.
13. Cavailles, V., et al., *Nuclear factor RIP140 modulates transcriptional activation by the estrogen receptor*. Embo J, 1995. **14**(15): p. 3741-51.
14. Carascossa, S., et al., *Receptor-interacting protein 140 is a repressor of the androgen receptor activity*. Mol Endocrinol, 2006. **20**(7): p. 1506-18.
15. Subramaniam, N., E. Treuter, and S. Okret, *Receptor interacting protein RIP140 inhibits both positive and negative gene regulation by glucocorticoids*. J Biol Chem, 1999. **274**(25): p. 18121-7.
16. Masuyama, H., et al., *Evidence for ligand-dependent intramolecular folding of the AF-2 domain in vitamin D receptor-activated transcription and coactivator interaction*. Mol Endocrinol, 1997. **11**(10): p. 1507-17.
17. Nephew, K.P., et al., *Studies of dehydroepiandrosterone (DHEA) with the human estrogen receptor in yeast*. Mol Cell Endocrinol, 1998. **143**(1-2): p. 133-42.

18. Sheeler, C.Q., M.W. Dudley, and S.A. Khan, *Environmental estrogens induce transcriptionally active estrogen receptor dimers in yeast: activity potentiated by the coactivator RIP140*. Environ Health Perspect, 2000. **108**(2): p. 97-103.
19. Henttu, P.M., E. Kalkhoven, and M.G. Parker, *AF-2 activity and recruitment of steroid receptor coactivator 1 to the estrogen receptor depend on a lysine residue conserved in nuclear receptors*. Mol Cell Biol, 1997. **17**(4): p. 1832-9.
20. Lee, C.H., C. Chinpaisal, and L.N. Wei, *Cloning and characterization of mouse RIP140, a corepressor for nuclear orphan receptor TR2*. Mol Cell Biol, 1998. **18**(11): p. 6745-55.
21. Christian, M., et al., *RIP140-targeted repression of gene expression in adipocytes*. Mol Cell Biol, 2005. **25**(21): p. 9383-91.
22. Anzick, S.L., et al., *AIB1, a steroid receptor coactivator amplified in breast and ovarian cancer*. Science, 1997. **277**(5328): p. 965-8.
23. Suen, C.S., et al., *A transcriptional coactivator, steroid receptor coactivator-3, selectively augments steroid receptor transcriptional activity*. J Biol Chem, 1998. **273**(42): p. 27645-53.
24. Azorsa, D.O., H.E. Cunliffe, and P.S. Meltzer, *Association of steroid receptor coactivator AIB1 with estrogen receptor-alpha in breast cancer cells*. Breast Cancer Res Treat, 2001. **70**(2): p. 89-101.
25. Louie, M.C., et al., *ACTR/AIB1 functions as an E2F1 coactivator to promote breast cancer cell proliferation and antiestrogen resistance*. Mol Cell Biol, 2004. **24**(12): p. 5157-71.
26. Carroll, J.S., et al., *Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1*. Cell, 2005. **122**(1): p. 33-43.
27. Carroll, J.S., et al., *Genome-wide analysis of estrogen receptor binding sites*. Nat Genet, 2006. **38**(11): p. 1289-97.
28. Charn, T.H., et al., *Genome-wide dynamics of chromatin binding of estrogen receptors alpha and beta: mutual restriction and competitive site selection*. Mol Endocrinol. **24**(1): p. 47-59.
29. Frasor, J., et al., *Selective estrogen receptor modulators: discrimination of agonistic versus antagonistic activities by gene expression profiling in breast cancer cells*. Cancer Res, 2004. **64**(4): p. 1522-33.
30. Kininis, M., et al., *Genomic analyses of transcription factor binding, histone acetylation, and gene expression reveal mechanistically distinct classes of estrogen-regulated promoters*. Mol Cell Biol, 2007. **27**(14): p. 5090-104.
31. Lin, C.Y., et al., *Whole-genome cartography of estrogen receptor alpha binding sites*. PLoS Genet, 2007. **3**(6): p. e87.
32. Liu, Y., et al., *The genome landscape of ERalpha- and ERbeta-binding DNA regions*. Proc Natl Acad Sci U S A, 2008. **105**(7): p. 2604-9.
33. Frasor, J., et al., *Gene expression preferentially regulated by tamoxifen in breast cancer cells and correlations with clinical outcome*. Cancer Res, 2006. **66**(14): p. 7334-40.
34. Manas, E.S., et al., *Structure-based design of estrogen receptor-beta selective ligands*. J Am Chem Soc, 2004. **126**(46): p. 15106-19.
35. Stauffer, S.R., et al., *Pyrazole ligands: structure-affinity/activity relationships and estrogen receptor-alpha-selective agonists*. J Med Chem, 2000. **43**(26): p. 4934-47.



36. Barnett, D.H., et al., *Estrogen receptor regulation of carbonic anhydrase XII through a distal enhancer in breast cancer*. Cancer Res, 2008. **68**(9): p. 3505-15.
37. Choi, I., et al., *Human estrogen receptor beta-specific monoclonal antibodies: characterization and use in studies of estrogen receptor beta protein expression in reproductive tissues*. Mol Cell Endocrinol, 2001. **181**(1-2): p. 139-50.
38. Zhang, Y., et al., *Model-based analysis of ChIP-Seq (MACS)*. Genome Biol, 2008. **9**(9): p. R137.
39. Huang da, W., B.T. Sherman, and R.A. Lempicki, *Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources*. Nat Protoc, 2009. **4**(1): p. 44-57.
40. Dennis, G., Jr., et al., *DAVID: Database for Annotation, Visualization, and Integrated Discovery*. Genome Biol, 2003. **4**(5): p. P3.
41. Malamas, M.S., et al., *Design and synthesis of aryl diphenolic azoles as potent and selective estrogen receptor-beta ligands*. J Med Chem, 2004. **47**(21): p. 5021-40.
42. Hartman, J., et al., *Estrogen receptor beta inhibits angiogenesis and growth of T47D breast cancer xenografts*. Cancer Res, 2006. **66**(23): p. 11207-13.
43. Bishop, J.M., *Molecular themes in oncogenesis*. Cell, 1991. **64**(2): p. 235-48.
44. Hunter, T., *Oncoprotein networks*. Cell, 1997. **88**(3): p. 333-46.
45. Perou, C.M., et al., *Molecular portraits of human breast tumours*. Nature, 2000. **406**(6797): p. 747-52.
46. Whitfield, M.L., et al., *Identification of genes periodically expressed in the human cell cycle and their expression in tumors*. Mol Biol Cell, 2002. **13**(6): p. 1977-2000.
47. Lupien, M., et al., *FoxA1 translates epigenetic signatures into enhancer-driven lineage-specific transcription*. Cell, 2008. **132**(6): p. 958-70.
48. Eeckhoute, J., et al., *Positive cross-regulatory loop ties GATA-3 to estrogen receptor alpha expression in breast cancer*. Cancer Res, 2007. **67**(13): p. 6477-83.
49. Albert, I., et al., *Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome*. Nature, 2007. **446**(7135): p. 572-6.
50. Lee, W., et al., *A high-resolution atlas of nucleosome occupancy in yeast*. Nat Genet, 2007. **39**(10): p. 1235-44.
51. Kaplan, N., et al., *The DNA-encoded nucleosome organization of a eukaryotic genome*. Nature, 2009. **458**(7236): p. 362-6.
52. Augereau, P., et al., *Transcriptional regulation of the human NR1P1/RIP140 gene by estrogen is modulated by dioxin signalling*. Mol Pharmacol, 2006. **69**(4): p. 1338-46.
53. Teyssier, C., et al., *Receptor-interacting protein 140 binds c-Jun and inhibits estradiol-induced activator protein-1 activity by reversing glucocorticoid receptor-interacting protein 1 effect*. Mol Endocrinol, 2003. **17**(2): p. 287-99.
54. Cowley, S.M., et al., *Estrogen receptors alpha and beta form heterodimers on DNA*. J Biol Chem, 1997. **272**(32): p. 19858-62.
55. Li, X., et al., *Single-chain estrogen receptors (ERs) reveal that the ERalpha/beta heterodimer emulates functions of the ERalpha dimer in genomic estrogen signaling pathways*. Mol Cell Biol, 2004. **24**(17): p. 7681-94.
56. Pace, P., et al., *Human estrogen receptor beta binds DNA in a manner similar to and dimerizes with estrogen receptor alpha*. J Biol Chem, 1997. **272**(41): p. 25832-8.
57. Pettersson, K., et al., *Mouse estrogen receptor beta forms estrogen response element-binding heterodimers with estrogen receptor alpha*. Mol Endocrinol, 1997. **11**(10): p. 1486-96.

58. Lanz, R.B., et al., *Global characterization of transcriptional impact of the SRC-3 coregulator*. Mol Endocrinol. **24**(4): p. 859-72.
59. Perissi, V., et al., *A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors*. Cell, 2004. **116**(4): p. 511-26.
60. Peterson, T.J., et al., *The silencing mediator of retinoic acid and thyroid hormone receptor (SMRT) corepressor is required for full estrogen receptor alpha transcriptional activity*. Mol Cell Biol, 2007. **27**(17): p. 5933-48.

## CHAPTER 5 CONCLUSION

In this thesis report, I have systematically mapped the genomic landscape of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 in multiple sub-lines of MCF-7 breast cancer cells, expressing various complements of ER $\alpha$  and ER $\beta$ , in the presence of E2. In addition to the genome-wide binding site cartographies, we also carried out gene expression profiling analyses to investigate the gene regulatory effects of ER $\alpha$  and ER $\beta$  in breast cancer cells. We were able to reveal several important ERs biological mechanisms in breast cancer by analyzing the global cartographies of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 together with microarray expression profiling data.

By analyzing ER $\alpha$  and ER $\beta$  chromatin binding sites, we found that ER $\alpha$  and ER $\beta$  bind to a similar, large number of sites in breast cancer cells containing only one ER subtype. However, in cells containing both ER $\alpha$  and ER $\beta$ , we observed that the two ERs appear to restrict each others chromatin binding and occupy fewer sites. This result suggests that, on a genome scale, there is active competition between the two receptors for chromatin binding sites when they are present together in cells, and they can no longer bind in an unobstructed manner to their native binding sites when they are present alone in the cells. Motifs enrichment analyses on ER $\alpha$  and ER $\beta$  binding sites suggest that, despite both ERs having highly conserved DNA-binding domain, there are differences in term of enriched motifs in the vicinity of ER $\alpha$  and ER $\beta$  binding sites. ER $\alpha$  binding sites are enriched in GATA and FOXA1 motifs, whereas ER $\beta$  sites are preferentially enriched in E2F motifs, suggesting that there are other cofactors such as FOXA1, GATA3, E2F, etc., enforcing the selectivity and range of ER $\alpha$  and ER $\beta$  binding.

Analysis of ER $\alpha$ , ER $\beta$ , SRC3, and RIP140 ChIP-seq data revealed that the coregulators SRC3 and RIP140 are recruited preferentially by both ER $\alpha$  and ER $\beta$  to the regulatory regions of estrogen-induced genes. Interestingly, SRC3 or RIP140 are seldom recruited by ER to estrogen-

repressed genes, which suggest that SRC3 and RIP140 are rarely used by ER for gene repression in breast cancer cells. Taken together, our study has provided new insight into the underlying mechanisms of ER-mediated transcriptional program, namely, that the SRC3-RIP140 complex might be playing a central role in the induction of ER target genes in breast cancer cells.

Gene ontology analysis of our gene chip microarray transcriptional profiling data enabled us to delineate a core set of genes that correlate with ER $\alpha$  proliferative and ER $\beta$  anti-proliferative effects in breast cancer cells. We found that ER $\beta$  activation by estradiol was associated with the inhibition of genes associated with cell proliferation and the up-regulation of pro-apoptotic genes and genes responding to DNA damage, whereas ER $\alpha$  activation was associated with the downregulation of pro-apoptotic genes and genes repressing transcription. These sets of genes may, at least in part, be responsible for the distinct roles of ER $\alpha$  and ER $\beta$  in breast cancer cell proliferation.

In summary, the data presented in this report has provided an integrated model in which the actions of cofactors such as FOXA1, GATA3, and E2F enforce the selectivity and range of ER $\alpha$  and ER $\beta$  binding and gene regulatory actions, with the coregulators SRC3 and RIP140 preferentially supporting the stimulatory actions of both receptors on gene expression in breast cancer cells. Furthermore, the genome-wide cartographies of ER $\alpha$  and ER $\beta$  and coregulators, SRC3 and RIP140, and the ER target genes provided in this thesis report will be invaluable for the investigation of ER signaling in breast cancer.